THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

November 12, 1996

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: *08/601,132*
FILING DATE: *February 14, 1996*

PRIORITY DOCUMENT

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

D. SMITH
Certifying Officer

PATENT APPLICATION SERIAL NO. 08601132

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FILE RECORD SHEET

SE18051 06/19/96 08601132 13-2853 150 101 38.00CR

PTO-1556
(5/87)

RECEPTOR LIGAND

This is a continuation-in-part of United States Patent Application Serial Number 08/585,895, filed January 12, 1996, which is a continuation-in-part of United States Patent Application Serial Number 08/510,133, filed August 1, 1995.

FIELD OF THE INVENTION

The present invention generally relates to the field of genetic engineering and more particularly to growth factors for endothelial cells and growth factor genes.

BACKGROUND OF THE INVENTION

Developmental growth, the remodelling and regeneration of adult tissues, as well as solid tumor growth, can only occur when accompanied by blood vessel formation. Angioblasts and hematopoietic precursor cells differentiate from the mesoderm and form the blood islands of the yolk sac and the primary vascular system of the embryo. The development of blood vessels from these early (*in situ*) differentiating endothelial cells is termed vasculogenesis. Major embryonic blood vessels are believed to arise via vasculogenesis, whereas the formation of the rest of the vascular tree is thought to occur as a result of vascular sprouting from pre-existing vessels, a process called angiogenesis, Risau, *et al.*, *Devel. Biol.*, *125*:441-450 (1988).

Endothelial cells give rise to several types of functionally and morphologically distinct vessels. When organs differentiate and begin to perform their specific functions, the phenotypic heterogeneity of endothelial cells increases. Upon angiogenic stimulation, endothelial cells may re-enter the cell cycle, migrate, withdraw from the cell cycle and subsequently differentiate again to form new vessels that are functionally adapted to their tissue environment. Endothelial cells undergoing angiogenesis degrade the underlying basement membrane and migrate, forming capillary sprouts that project into the perivascular stroma. Ausprunk, *et al.*, *Microvasc. Rev.*, *14*:51-65 (1977). Angiogenesis during tissue development and regeneration depends on the tightly controlled processes of endothelial cell proliferation, migration, differentiation, and survival. Dysfunction of the endothelial cell regulatory system is a key feature of many diseases. Most significantly, tumor

growth and metastasis have been shown to be angiogenesis dependent. Folkman, *et al.*, *J. Biol. Chem.*, *267*:10931-10934 (1992).

Key signals regulating cell growth and differentiation are mediated by polypeptide growth factors and their transmembrane receptors, many of which are tyrosine kinases. Autophosphorylated peptides within the tyrosine kinase insert and carboxyl-terminal sequences of activated receptors are commonly recognized by kinase substrates involved in signal transduction for the readjustment of gene expression in responding cells. Several families of receptor tyrosine kinases have been characterized. Van der Geer, *et al.*, *Ann. Rev. Cell Biol.*, *10*:251-337 (1994). The major growth factors and receptors transducing angiogenic stimuli are schematically shown in Figure 1.

Fibroblast growth factors are also known to be involved in the regulation of angiogenesis. They have been shown to be mitogenic and chemotactic for cultured endothelial cells. Fibroblast growth factors also stimulate the production of proteases, such as collagenases and plasminogen activators, and induce tube formation by endothelial cells. Saksela, *et al.*, *Ann. Rev. Cell Biol.*, *4*:93-126 (1988). There are two general classes of fibroblast growth factors, FGF-1 and FGF-2, both of which lack conventional signal peptides. Both types have an affinity for heparin and FGF-2 is bound to heparin sulfate proteoglycans in the subendothelial extracellular matrix from which it may be released after injury. Heparin potentiates the stimulation of endothelial cell proliferation by angiogenic FGFs, both by protecting against denaturation and degradation and dimerizing the FGFs. Cultured endothelial cells express the FGF-1 receptor but no significant levels of other high-affinity fibroblast growth factor receptors.

Among other ligands for receptor tyrosine kinases, the platelet derived growth factor, PDGF-BB, has been shown to be weakly angiogenic in the chick chorioallantoic membrane. Risau, *et al.*, *Growth Factors*, *7*:261-265 (1992). Transforming growth factor α (TGFα) is an angiogenic factor secreted by several tumor cell types and by macrophages. Hepatocyte growth factor (HGF), the ligand of the c-*met* proto-oncogene-encoded receptor, also is strongly angiogenic.

Recent evidence shows that there are endothelial cell specific growth factors and receptors that may be primarily responsible for the stimulation of endothelial cell growth, differentiation and certain differentiated

functions. The best studied of these is vascular endothelial growth factor (VEGF), a member of the PDGF family. Vascular endothelial growth factor is a dimeric glycoprotein of disulfide-linked 23 kDa subunits. Other reported effects of VEGF include the mobilization of intracellular calcium, the induction of plasminogen activator and plasminogen activator inhibitor-1 synthesis, stimulation of hexose transport in endothelial cells, and promotion of monocyte migration *in vitro*. Four VEGF isoforms, encoded by distinct mRNA splice variants, appear to be equally capable of stimulating mitogenesis in endothelial cells. However, each isoform has a different affinity for cell surface proteoglycans, which behave as low affinity receptors for VEGF. The 121 and 165 amino acid isoforms of VEGF (VEGF121 and VEGF165) are secreted in a soluble form, whereas the isoforms of 189 and 206 amino acid residues remain cell surface associated and have a strong affinity for heparin.

VEGF was originally purified from several sources on the basis of its mitogenic activity toward endothelial cells, and also by its ability to induce microvascular permeability, hence it is also called vascular permeability factor (VPF). VEGF produces signals through two receptor tyrosine kinases, VEGFR-1 (FLT-1) and VEGFR-2 (KDR/Flk-1), which are expressed specifically on endothelial cells. The VEGF-related placenta growth factor (PlGF) was recently shown to bind to VEGFR-1 with high affinity. PlGF was able to enhance the growth factor activity of VEGF, but it did not stimulate endothelial cells on its own. Naturally occurring VEGF/PlGF heterodimers were nearly as potent mitogens as VEGF homodimers for endothelial cells.

The pattern of VEGF expression suggests its involvement in the development and maintenance of the normal vascular system and in tumor angiogenesis. During murine development, the entire 7.5 day post-coital (p.c.) endoderm expresses VEGF and the ventricular neuroectoderm produces VEGF at the capillary ingrowth stage. See Breier, *et al.*, *Development*, *114*:521-523 (1992). On day two of quail development, the vascularized area of the yolk sac as well as the whole embryo show expression of VEGF. In addition, epithelial cells next to fenestrated endothelia in adult mice show persistent VEGF expression, suggesting a role in the maintenance of this specific endothelial phenotype and function.

Two high affinity receptors for VEGF have been characterized. These are VEGFR-1/Flt-1 (fms-like tyrosine kinase-1) and VEGFR-2/Kdr/Flk-

1 (kinase insert domain containing receptor/fetal liver kinase-1). Those receptors are classified in the PDGF-receptor family, but they have seven rather than five immunoglobulin-like loops in their extracellular domain and they possess a longer kinase insert than normally observed in this family. The expression of VEGF receptors occurs mainly in vascular endothelial cells, although some may be present on monocytes and melanoma cells. Only endothelial cells have been reported to proliferate in response to VEGF, and endothelial cells from different sources show different responses. Thus, the signals mediated through VEGFR-1 and VEGFR-2 appear to be cell type specific.

The Flt4 receptor tyrosine kinase (VEGFR-3) is closely related in structure to the products of the VEGFR-1 and VEGFR-2 genes. Despite this similarity, the mature form of Flt4 differs from the VEGF receptors in that it is proteolytically cleaved in the extracellular domain into two disulfide-linked polypeptides. Pajusola *et al.*, *Cancer Res.*, *52*:5738-5743 (1992). The 4.5 and 5.8 kb Flt4 mRNAs encode polypeptides which differ in their C-termini due to the use of alternative 3' exons. The VEGFs do not show specific binding to Flt4 or induce its autophosphorylation.

Expression of Flt4 appears to be more restricted than expression of VEGFR-1 or VEGFR-2. The expression of Flt4 first becomes detectable by *in situ* hybridization in the angioblasts of head mesenchyme, the cardinal vein, and extraembryonically in the allantois of 8.5 day p.c. mouse embryos. In 12.5 day p.c. embryos the Flt4 signal is observed in developing venous and presumptive lymphatic endothelia, but arterial endothelia appear negative. During later stages of development, Flt4 mRNA becomes restricted to developing lymphatic vessels. Only the lymphatic endothelia and some high endothelial venules express Flt4 mRNA in adult human tissues and increased expression occurs in lymphatic sinuses in metastatic lymph nodes and in lymphangioma. These results support the theory of the venous origin of lymphatic vessels.

Five endothelial cell specific receptor tyrosine kinases, Flt-1 (VEGFR-1), KDR/Flk-1 (VEGFR-2), Flt4, Tie and Tek/Tie-2 have so far been described, which possess the intrinsic tyrosine kinase activity essential for signal transduction. Targeted mutations inactivating Flt-1, Flk-1, Tie and Tek in mouse embryos have indicated their essential and specific roles in

vasculogenesis and angiogenesis at the molecular level. VEGFR-1 and VEGFR-2 bind VEGF with high affinity (Kd 16 pM and 760 pM, respectively) and VEGFR-1 also binds the related placenta growth factor (PlGF; Kd about 200 pM), while the ligands for Tie, Tek, and Flt4 have not heretofore been reported.

SUMMARY OF THE INVENTION

The present invention provides a ligand for the Flt4 receptor tyrosine kinase. Thus, the invention provides a purified and isolated polypeptide which specifically binds to the Flt4 receptor tyrosine kinase. In a preferred embodiment, the ligand comprises a fragment of the amino acid sequence shown in SEQ ID NO: 33 which specifically binds to the human Flt4 receptor tyrosine kinase.

The present invention also provides one or more precursors of an Flt4 ligand, wherein one such precursor (designated "prepro-VEGF-C") comprises the complete amino acid sequence (amino acid residues -102 to 317) shown in SEQ ID NO: 33. Thus, the invention includes a purified and isolated polypeptide having the amino acid sequence of residues -102 to 317 shown in SEQ ID NO: 33.

A putative 102 amino acid leader (prepro) peptide has been identified in the amino acid sequence shown in SEQ ID NO: 33. Thus, in a related aspect, the invention includes a purified and isolated polypeptide having the amino acids sequence of residues 1-317 shown in SEQ ID NO: 33.

The expressed Flt4 ligand precursor is proteolytically cleaved upon expression to produce an approximately 23 kD peptide which is the Flt4 ligand (herein designated VEGF-C). Thus, in a preferred embodiment of the invention, an Flt4 ligand is provided which is the cleavage product of the precursor peptide shown in SEQ ID NO: 33 and which has a molecular weight of approximately 23 kD under reducing conditions. A putative VEGF-C precursor, having an observed molecular weight of about 32 kD, also is considered an aspect of the invention.

From the foregoing, it will be apparent that an aspect of the invention includes a fragment of the purified and isolated polypeptide having the amino acid sequence of residues -102 to 317 shown in SEQ ID NO: 33, the fragment being capable of specifically binding to Flt4 receptor tyrosine

kinase. In a preferred embodiment, the invention includes a fragment having an apparent molecular weight of approximately 23 kD as assessed by SDS-PAGE under reducing conditions.

In a related aspect, the invention includes a purified and isolated polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, the polypeptide comprising a fragment of the purified and isolated polypeptide according to claim 2, the fragment being capable of specifically binding to Flt4 receptor tyrosine kinase. Similarly, the invention includes a polypeptide having an amino acid sequence comprising a portion of SEQ ID NO:2, the portion encoding a fragment capable of specifically binding to Flt4.

Evidence suggests that the amino acids essential for retaining Flt4 ligand activity are contained within approximately amino acids 1-115 of SEQ ID NO: 33, and that a carboxy-terminal proteolytic cleavage to produce a mature, naturally-occurring Flt4 ligand occurs within approximately amino acids 115-180 of SEQ ID NO: 33. Accordingly, preferred polypeptides of the invention include polypeptides comprising amino acids 1-115, 1-116, 1-117, 1-118, 1-119, 1-120, 1-121, 1-122, 1-123, 1-124 ... 1-178, 1-179, and 1-180 of SEQ ID NO: 33, wherein the polypeptides specifically bind to an Flt4 receptor tyrosine kinase. A preferred Flt4 ligand comprises approximately amino acids 1-115 of SEQ ID NO: 33. Another preferred polypeptide of the invention comprises approximately amino acids 1-180 of SEQ ID NO: 33.

The present invention also provides a cDNA encoding a novel polypeptide, designated VEGF-C, that is structurally homologous to VEGF. VEGF-C is a ligand for the FLT4 receptor tyrosine kinase (VEGFR-3), a receptor tyrosine kinase related to VEGFR-1 and VEGFR-2 that does not bind VEGF. VEGFR-3 is expressed in venous and lymphatic endothelia of fetal tissues and predominantly in lymphatic endothelial of adult tissues. Kaipainen et al., *Cancer Res.*, *54*:6571-77 (1994); Kaipainen et al., *Proc. Natl. Acad. Sci. USA*, *92*:3566-70 (1995).

Thus, aspects of the invention include purified and isolated nucleic acids encoding polypeptides and polypeptide fragments of the invention, vectors which comprise nucleic acids of the invention; and host cells transformed or transfected with nucleic acids or vectors of the invention. For example, in a preferred embodiment, the invention includes a purified and isolated nucleic acid (e.g., a DNA or an RNA) encoding an Flt4 ligand

precursor. Due to the degeneracy of the genetic code, numerous such coding sequences are possible, each having in common the coding of the amino acid sequence shown in SEQ ID NO: 33. As set forth above, the invention includes polypeptides which comprise a portion of the amino acid sequence shown in SEQ ID NO: 33 and which bind the Flt4 receptor tyrosine kinase (herein designated VEGFR-3); the invention also is intended to include nucleic acids encoding these polypeptides. Ligand precursors according to the invention, when expressed in an appropriate host cell, produce, via cleavage, a peptide which binds specifically to the Flt4 receptor tyrosine kinase (VEGFR-3). The nucleotide sequence shown in SEQ ID NO:32 contains a preferred nucleotide sequence encoding the Flt4 ligand (VEGF-C).

The present invention also provides a cell line which produces an Flt4 ligand. In a preferred embodiment, the ligand may be purified and isolated directly from the cell culture medium. Also provided are vectors comprising a DNA encoding the Flt4 ligand, and host cells comprising the vectors. Preferred vectors of the invention are expression vectors wherein nucleic acids of the invention are operatively connected to appropriate promoters and other control sequences, such that appropriate host cells transfromed or transfected with the vectors are capable of expressing the Flt4 ligand. A preferred vector of the invention is plasmid pFLT4-L, having ATCC accession no. 97321

The invention further includes a method of making polypeptides of the invention. In a preferred method, a nucleic acid or vector of the invention is expressed in a host cell, and a polypeptide of the invention is purified from the host cell or the host cell's growth medium.

In a related embodiment, the invention includes a method of making a polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, comprising the steps of: (a) transforming or transfecting a host cell with a nucleic acid of the invention; (b) cultivating the host cell to express the nucleic acid; and (c) purifying a polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase from the host cell or from the host cell's growth media.

The invention also is intended to include purified and isolated polypeptide ligands of Flt4 produced by methods of the invention.

In another aspect, the invention includes an antibody which is specifically reactive with polypeptides of the invention. Antibodies, both monoclonal and polyclonal, may be made against a ligand of the invention according to standard techniques in the art. Such antibodies may be used in diagnostic applications to monitor angiogenesis, vascularization, lymphatic vessels and their disease states, wound healing, or certain hematopoietic or leukemia cells, or they may be used to block or activate the Flt4 receptor.

Ligands according to the invention may be labeled with a detectable label and used to identify their corresponding receptors *in situ*. Labeled Flt4 ligand and anti-Flt4 ligand antibodies may be used as imaging agents in the detection of lymphatic vessels, high endothelial venules, and Flt4 receptors expressed in histochemical tissue sections. The ligand or antibody may be covalently or non-covalently coupled to a suitable supermagnetic, paramagnetic, electron dense, echogenic, or radioactive agent for imaging. Other, non-radioactive labels, such as biotin and avidin, may also be used.

The present invention also provides diagnostic and clinical applications for claimed ligands. In a preferred embodiment, Flt4 ligands or precursors are used to accelerate angiogenesis, *e.g.*, during wound healing, or to promote the endothelial functions of lymphatic vessels. A utility for VEGF-C is suggested as an inducer of angiogenesis also in tissue transplantation, in eye diseases, in the formation of collateral vessels to around arterial stenoses and into injured tissues after infarction. Ligands may be applied in any suitable manner using an appropriate pharmaceutically-acceptable vehicle, e.g., a pharmaceutically-acceptable diluent, adjuvent, or carrier. Ligands also may be used to quantify future metastatic risk by assaying biopsy material for the presence of active receptors or ligands in a binding assay or kit using detectably-labeled ligand. An Flt4 ligand according to the invention also may be used to promote re-growth or permeability of lymphatic vessels in, for example, organ transplant patients. Ligands according to the invention also may be used to treat or prevent inflammation, edema, aplasia of the lymphatic vessels, lymphatic obstruction, elephantiasis, and Milroy's disease. Finally, Flt4 ligands may be used to stimulate lymphocyte production and maturation, and to promote or inhibit trafficking of leukocytes between tissues and lymphatic vessels or to affect migration in and out of the thymus.

Inhibitors of the Flt4 ligand may be used to control endothelial cell proliferation and lymphangiomas. For example, such inhibitors may be used to arrest metastatic growth or spread, or to control other aspects of endothelial cell expression and growth. Inhibitors include antibodies, antisense oligonucleotides, and peptides which block the Flt4 receptor, all of which are intended as aspects of the invention.

Polynucleotides of the invention such as those described above, fragments of those polynucleotides, and variants of those polynucleotides with sufficient similarity to the non-coding strand of those polynucleotides to hybridize thereto under stringent conditions all are useful for identifying, purifying, and isolating polynucleotides encoding other (non-human) mammalian forms of VEGF-C. Thus, such polynucleotide fragments and variants are intended as aspects of the invention. Exemplary stringent hybridization conditions are as follows: hybridization at 42 C in 5X SSC, 20 mM $NaPO_4$, pH 6.8, 50% formamide; and washing at 42 C in 0.2X SSC. Those skilled in the art understand that it is desirable to vary these conditions empirically based on the length and the GC nucleotide base content of the sequences to be hybridized, and that formulas for determining such variation exist. See, e.g., Sambrook et al., *Molecular Cloning: a Laboratory Manual*, Second Edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory (1989).

Moreover, purified and isolated polynucleotides encoding other (non-human) mammalian VEGF-C forms also are aspects of the invention, as are the polypeptides encoded thereby, and antibodies that are specifically immunoreactive with the non-human VEGF-C variants. Thus, the invention includes a purified and isolated mammalian VEGF-C polypeptide, and also a purified and isolated polynucleotide encoding such a polypeptide.

In one embodiment, the invention includes a purified and isolated polypeptide having the amino acid sequence of residues 1 to 415 of SEQ ID NO: 41, which sequence corresponds to a putative mouse VEGF-C precursor. The putative mouse VEGF-C precursor is believed to be processed into a mature mouse VEGF-C in a manner analogous to the processing of the human prepro-polypeptide. Thus, in a related aspect, the invention includes a purified and isolated polypeptide capable of specifically binding to an Flt4 receptor tyrosine kinase (e.g., a human or mouse Flt-4 receptor tyrosine

kinase), the polypeptide comprising a fragment of the purified and isolated polypeptide having the amino acid sequence of residues 1 to 415 of SEQ ID NO: 41, the fragment being capable of specifically binding to the Flt4 receptor tyrosine kinase. The invention further includes purified and isolated nucleic acids encoding the foregoing polypeptides, such as a nucleic acid comprising all or a portion of the sequence shown in SEQ ID NO: 40.

BRIEF DESCRIPTION OF THE DRAWING

Figure 1 is a schematic diagram showing major endothelial cell receptor tyrosine kinases and growth factors involved in vasculogenesis and angiogenesis.

Figure 2 schematically depicts the construction of the pLTRFlt4i expression vector.

Figure 3 schematically depicts the construction of the baculovirus vector encoding a secreted soluble Flt4 extracellular domain (Flt4EC).

Figure 4 shows results of stimulation of Flt4 autophosphorylation by conditioned medium from PC-3 cell cultures.

Figures 5A, 5B, and 5C show that the major tyrosyl phosphorylated polypeptide of Flt4-transfected cells stimulated with PC-3 conditioned medium is the 125 kD Flt4 polypeptide (VEGFR-3), and also that the Flt4 stimulating activity is not adsorbed to heparin-sepharose.

Figure 6 shows Western immunoblotting analysis of the Flt4 ligand activity isolated from PC-3 conditioned medium.

Figure 7 shows results of gel electrophoresis of fractions from the Western analysis of Flt4 ligand (VEGF-C) isolated from PC-3 conditioned medium.

Figure 8 shows results of Western analysis of Flt4 autophosphorylation induced by either the Flt4 ligand (VEGF-C), VEGF, or PlGF.

Figure 9A schematically depicts the cloning and analysis of the Flt4 ligand, VEGF-C. The VEGF homologous region (dark shaded box) and amino and carboxyl terminal propeptides (light shaded and unshaded boxes, respectively) as well as putative signal sequence (ss) are depicted between 5' and 3' untranslated (ut) nucleic acid regions. The cleavage sites for the signal

sequence and the amino and carboxyl terminal propeptides are indicated with triangles.

Figure 9B shows the nucleotide and deduced amino acid sequence of a Flt4 ligand cDNA (without adaptor and poly-A sequences). The cleavage site for the putative amino terminal prepro leader sequence is indicated with a shaded triangle.

Figure 10 shows a comparison of the deduced amino acid sequences of PDGF-A, -B, PlGF-1, VEGF 3[illegible]7, four VEGF isoforms, and Flt4 ligand (VEGF-C).

Figure 11 shows the stimulation of autophosphorylation of the Flt4 receptor by conditioned medium from cells transfected with the pREP7 expression vector containing the VEGF-C-encoding cDNA insert of plasmid FLT4-L.

Figure 12 shows Northern blotting analysis of the genes encoding VEGF, VEGF-B, AND VEGF-C (indicated by "FLT4-L") in two human tumor cell lines.

Figure 13A is an autoradiograph showing recombinant VEFG-C isolated following a pulse-chase experiment and electrophoresed via SDS-PAGE under reducing conditions.

Figure 13B is a photograph of polyacrylamide gel showing that recombinant VEFG-C forms are disulfide-linked in nonreducing conditions.

Figure 14A and 14B depict Western blots showing that VEGF-C stimulates autophosphorylation of VEGFR-2 (KDR) but has no effect on PDGFR-β phosphorylation.

Figure 15A and 15B show that VEGF-C stimulates endothelial cell migration in a three-dimensional collagen gel assay.

Figure 16A shows the expression of VEGF-C mRNA in human adult tissues.

Figure 16B shows the expression of VEGF, VEGF-B, and VEGF-C in selected human fetal tissues.

Figure 17 schematically depicts the chromosomal localization of the VEGF-C gene.

Figure 18 is a Northern blot hybridization study showing the effects of hypoxia on the mRNA expression of VEGF (VEGF-A), VEGF-B and VEGF-C.

Figure 19 depicts autoradiograms from a pulse-chase immunoprecipitation experiment wherein cells transfected with a VEGF-C expression vector (VEGF-C) and mock transfected cells (M) were pulse-labeled with radioactive amino acids and chased for varying lengths of time.

DETAILED DESCRIPTION OF THE INVENTION

Described herein is the isolation of a novel vascular endothelial growth factor and the cloning of a DNA encoding this novel growth factor from a cDNA library prepared from the human prostatic adenocarcinoma cell line PC-3. The isolated cDNA encodes a protein which is proteolytically processed and secreted to cell culture medium. The secreted protein, designated VEGF-C, binds to the extracellular domain of Flt4 and induces tyrosine autophosphorylation of Flt4 and VEGFR-2. VEGF-C also stimulates the migration of endothelial cells in collagen gel.

The present invention also is directed to novel growth factor polypeptides which are ligands for the Flt4 receptor tyrosine kinase (VEGFR-3). Ligands of the invention are members of a family of platelet-derived growth factors/vascular endothelial growth factors which promote mitosis and proliferation of vascular endothelial cells and/or mesodermal cells. As described in greater detail in Examples 4 and 5, ligands recognizing the Flt4 receptor tyrosine kinase were purified from a PC-3 prostatic adenocarcinoma cell line (ATCC CRL1435). When applied to a population of cells expressing the Flt4 receptor, ligands of the invention stimulate autophosphorylation, resulting in receptor activation.

A ligand according to the invention may be expressed as a larger precursor which is cleaved to produce the ligand. A coexpressed region in some cases results from alternative splicing of RNA of the ligand gene. Such a co-expressed region may be a function of the particular expression system used to obtain the ligand. The skilled artisan understands that in recombinant production of proteins, additional sequence may be expressed along with a functional peptide depending upon the particular recombinant construct used to express the protein, and subsequently removed to obtain the desired ligand. In some cases the recombinant ligand can be made lacking certain residues of the endogenous/natural ligand. Moreover, it is well-known

in that conservative replacements may be made in a protein which do not alter the function of the protein. Accordingly, it is anticipated that such alterations are within the scope of the claims. Moreover, it is anticipated that one or more VEGF-C precursors (the largest putative native VEGF-C precursor having the complete amino acid sequence from residue -102 to residue 317 of SEQ ID NO: 33) is capable of stimulating the Flt4 ligand without any further processing, in a manner similar to that in which VEGF stimulates its receptor in its unprocessed form after the secretion and concomitant release of the signal sequence.

Results reported herein show that Flt4 (VEGFR-3) transmits signals for the VEGF-C novel growth factor. This conclusion is based on the specific binding of VEGF-C to recombinant Flt4EC (Flt4 extracellular domain) protein and the induction of VEGFR-3 autophosphorylation by medium from VEGF-C transfected cells. In contrast, neither VEGF nor PlGF showed specific binding to VEGFR-3 or induced its autophosphorylation.

As set forth in greater detail below, the putative prepro-VEGF-C has a deduced molecular mass of 46,883; a putative prepro-VEGF-C processing intermediate has an observed molecular weight of about 32kD; and mature VEGF-C isolated form conditioned media has a molecular weight of about 23kD as assessed by SDS-Page under reducing conditions. A major part of the difference in the observed molecular mass of the purified and recombinant VEGF-C and the deduced molecular mass of the prepro-VEGF-C encoded by the VEGF-C open reading frame (ORF) is attributable to proteolytic removal of sequences at the amino-terminal and carboxyl-terminal regions of the prepro-VEGF-C polypeptide. However, proteolytic cleavage of the putative 102 amino acid prepro- leader sequence is not believed to account for the entire difference between the deduced molecular mass of 46,883 and the observed mass of about 23kD, because the deduced molecular weight of a polypeptide consisting of amino acids 1-317 of SEQ ID NO:33 is 35,724 kD. It is believed that a portion of the observed difference in molecular weights is attributable to proteolytic removal of amino acid residues in the carboxyl terminal region of the VEGF-C precursor. By extrapolation from studies of the structure of PDGF (Heldin, *et al.*, *Growth Factors*, 8:245-52 (1993)), one can speculate that the region critical for receptor binding and activation by VEGF-C is contained within the amino-terminal first 180 or so amino acid

residues the secreted of VEGF-C protein lacking the putative prepro leader sequence. In fact, the region critical for receptor binding and activation by VEGF-C is believed to be contained within the first approximately 120 amino acid residues of the secreted VEGF-C protein lacking the prepro leader sequence. Thus, the 23 kD polypeptide binding VEGFR-3 is likely to represent the VEGF-homologous domain. After biosynthesis, the nascent VEGF-C polypeptide may be glycosylated at three putative N-linked glycosylation sites identified in the deduced VEGF-C amino acid sequence. Polypeptides containing modifications, such as N-linked glycosylations, are intended as aspects of the invention.

The carboxyl terminal amino acid sequences, which increase the length of the VEGF-C polypeptide in comparison with other ligands of this family, show a pattern of spacing of cysteine residues reminiscent of the Balbaini ring 3 protein (BR3P) sequence (Dignam and Case, *Gene, 88*:133-40 (1990); Paulsson, *et al., J. Mol. Biol., 211*:331-49 (1990)). This novel C-terminal silk protein-like structural motif of VEGF-C may fold into an independent domain, which, on the basis of the considerations above, is at least partially cleaved off after biosynthesis. Interestingly, at least one cysteine motif of the BR3P type is also found in the carboxyl terminus of VEGF. In our experiments both the putative precursor and cleaved ligand were detected in the cell culture media, although processing was apparently cell-associated on the basis of the pulse-chase experiments. The determination of the amino-terminal and carboxy-terminal sequences of VEGF-C isolates will allow the identification of the proteolytic processing sites. The generation of antibodies against different parts of the pro-VEGF-C molecule will allow the exact determination of the precursor-product relationship and ratio, their cellular distribution, and the kinetics of processing and secretion.

VEGF-C has a conserved pattern of eight cysteine residues, which may participate in the formation of intra- and interchain disulfide bonds, creating an antiparallel dimeric biologically active molecule, similar to PDGF. Mutational analysis of the cysteine residues involved in the interchain disulfide bridges have shown that, in contrast to PDGF, VEGF dimers need to be held together by these covalent interactions in order to maintain biological activity. Disulfide linking of the VEGF-C polypeptide chain was evident in the analysis of VEGF-C in nonreducing conditions.

VEGFR-3, which distinguishes between VEGF and VEGF-C, is closely related to structure to VEGFR-1 and VEGFR-2. Finnerty, *et al.*, *Oncogene*, *8*:2293-98 (1993); Galland, *et al.*, *Oncogene*, *8*:1233-40 (1993); Pajusola, *et al.*, *Cancer Res.*, *52*:5738-43 (1992). However, the mature form of VEGFR-3 differs from the two other VEGFRs in that it is proteolytically cleaved in the extracellular domain into two disulfide-linked polypeptides. Pajusola, *et al.*, *Oncogene*, *9*:3545-55 (1994). Another difference is that the 4.5 and 5.8 kb VEGFR-3 mRNAs encode polypeptides differing in their C-termini and apparently in their signalling properties due to the use of alternative 3' exons. Borg et al., *Oncogene*, *10*:973-84 (1995); Pajusola et al., *Oncogene*, *8*:2931-37 (1993).

Besides VEGFR-3, VEGFR-2 tyrosine kinase also is activated in response to VEGF-C. VEGFR-2 mediated signals cause striking changes in the morphology, actin reorganization and membrane ruffling of porcine aortic endothelial cells overexpressing this receptor. In these cells, VEGFR-2 also mediated ligand-induced chemotaxis and mitogenicity. Waltenberger et al., *J. Biol. Chem*, *269*:26988-95 (1994). Similarly, the receptor chimera CSF-1R/VEGFR-3 was mitogenic when ectopically expressed in NIH3T3 fibroblastic cells, but not in porcine aortic endothelial cells (Pajusola et al., 1994). Consistent with such results, the bovine capillary endothelial (BCE) cells, which express VEGFR-2 mRNA but very little or no VEGFR-1 or VEGFR-3 mRNAs, showed enhanced migration when stimulated with VEGF-C. Light microscopy of the BCE cell cultures in collagen gel also suggested that VEGF-C stimulated the proliferation of these cells. The already existing data thus indicate that the VEGF ligands and receptors show a great specificity in their signalling, which may be cell type dependent.

The expression pattern of the VEGFR-3 (Kaipainen et al., *Proc. Natl. Acad. Sci. USA*, *92*:3566-70 (1995)) suggests that VEGF-C may function in the formation of the venous and lymphatic vascular systems during embryogenesis. Constitutive expression of VEGF-C in adult tissues shown herein further suggests that this gene product also is involved in the maintenance of the differentiated functions of the lymphatic endothelium where VEGFR-3 is expressed (Kaipainen et al., 1995). Lymphatic capillaries do not have well formed basal laminae and an interesting possibility remains that the silk-like BR3P motif is involved in producing a supramolecular structure which

could regulate the availability of VEGF-C in tissues. However, as shown here, VEGF-C also activates VEGFR-2, which is abundant in proliferating endothelial cells of vascular sprouts and branching vessels of embryonic tissues, but decreased in adult tissues. Millauer et al., *Nature*, *367*:576-78 (1993). These data have suggested that VEGFR-2 is a major regulator of vasculogenesis and angiogenesis. VEGF-C may thus have a unique effect in lymphatic endothelium and a more redundant function shared with VEGF in angiogenesis and possibly permeability regulation of several types of endothelia. Because VEGF-C stimulates the VEGFR-2 and promotes endothelial migration, a utility for VEGF-C is suggested as an inducer of angiogenesis of blood and lymphatic vessels in wound healing, tissue transplantation, in eye diseases, in the formation of collateral vessels to around arterial stenoses and into injured tissues after infarction.

Taken together, these results show an increased complexity of signalling in the vascular endothelium. They reinforce the concept that when organs differentiate and begin to perform their specific functions, the phenotypic heterogeneity of endothelial cells increases in several types of functionally and morphologically distinct vessels. However, upon suitable angiogenic stimuli, endothelial cells can re-enter the cell cycle, migrate, withdraw from the cell cycle and subsequently differentiate again to form new vessels that are functionally adapted to their tissue environment. This process of angiogenesis concurrent with tissue development and regeneration depends on the tightly controlled balance between positive and negative signals for endothelial cell proliferation, migration, differentiation and survival. Previously-identified growth factors promoting angiogenesis include the fibroblast growth factors, hepatocyte growth factor/scatter factor, PDGF and TGF-α. (See, e.g., Folkman, *Nature Med.* *1*:27-31 (1995); Friesel and Maciag, *FASEB J.* *9*:919-25 (1995); Mustonen and Alitalo, *J. Cell Biol.*, *129*:895-98 (1995). However, VEGF has been the only growth factor relatively specific for endothelial cells. The newly identified factors VEGF-B and VEGF-C thus increase our understanding of the complexity of the specific and redundant positive signals for endothelial cells involved in vasculogenesis, angiogenesis, permeability and perhaps also other endothelial functions.

Also described herein is the localization of the VEGF-C genes in human chromosomes by analysis of somatic cell hybrids and fluorescence *in*

situ hybridization (FISH). Southern blotting and polymerase chain reaction analysis of somatic cell hybrids and fluorescence *in situ* hybridization of metaphase chromosomes was used to assess the chromosomal localization of the VEGF-C gene. The VEGF-C gene was located on chromosome 4q34, close to the human aspartylglucosaminidase gene previously mapped to 4q34-35. The VEGF-C locus in 4q34 is a candidate target for mutations leading to vascular malformations or cardiovascular diseases. Expression studies by Northern blotting and hybridization show abundant VEGF-C expression in heart and skeletal muscle; other tissues, such as lung and kidney, also express this gene. Whereas PlGF is predominantly expressed in the placenta, the expression patterns of the VEGFs overlap in many tissues, which suggests that they may form heterodimers and interact to exert their physiological functions.

Targeted mutagenesis leading to inactivation of the VEGF receptor loci in the mouse genome have shown that VEGFR-1 is necessary for the proper organization of endothelial cells forming the vascular endothelium, while VEGFR-2 is necessary for the generation of both endothelial and hematopoietic cells. This suggests that the four genes of the VEGF family can be targets for mutations leading to vascular malformations or cardiovascular diseases.

The following Examples illustrate preferred embodiments of the invention, wherein the isolation, characterization, and function of Flt4 ligands and ligand-encoding nucleic acids according to the invention are shown.

EXAMPLE 1

Production of pLTRFlt4l expression vector

Construction of the LTR-Flt4l vector is schematically shown in Figure 2. The full-length Flt4s cDNA (Genbank Accession No. X68203, SEQ ID NO: 36) was assembled by first subcloning the S2.5 fragment, reported in Pajusola *et al.*, *Cancer Res.* 52:5738-5743 (1992), incorporated by reference herein, containing base pairs 56-2534 of the Flt4s into the *EcoRI* site of the pSP73 vector (Promega, Madison, WI).

Since cDNA libraries used for screening of Flt4 cDNAs did not contain its most 5' protein-coding sequences, inverse PCR was used for the amplification of the 5' end of Flt4 corresponding to the first 12 amino acid

residues (MQRGAALCLRLW). Poly(A)$^+$ RNA was isolated from the human HEL erythroleukemia cells and double-stranded cDNA copy was synthesized using the Amersham cDNA Synthesis System Plus kit and a gene specific primer: 5'-TGTCCTCGCTGTCCTTGTCT-3' (SEQ ID NO: 1), which was located 195 bp downstream of the 5' end of clone S2.5. Double stranded cDNA was treated with T4 DNA polymerase to blunt the ends and cDNA was purified with Centricon 100 (Amicon Inc., Beverly, MA). Circularization of the blunt-ended cDNA was performed in a total volume of 150 microliters. The reaction mixture contained ligation buffer, 5% PEG-8000, 1 mM DTT and 8U of T4 DNA ligase (New England Biolabs). Ligation was carried out at 16°C for 16 hours. Fifteen microliters of this reaction mix were used in a standard 100 microliters PCR reaction containing 100 ng of specific primers including *SacI* and *PstI* restriction sites, present in this segment of the Flt4 cDNA, and 1 unit of Taq DNA polymerase (Perkin Elmer Cetus). Two rounds of PCR were performed using 33 cycles (denaturation at 95°C for 1 minute, annealing at 55°C for 2 minutes and elongation at 72°C for 4 minutes). The PCR mixture was treated sequentially with the *SacI* and *PstI* restriction enzymes, and after purification with MagicPCR Preps (Promega), DNA fragments were subcloned into the pGEM3Zf(+) vector for sequencing. The sequence obtained corresponds to the 5' end of the Flt4s cDNA clone deposited in the Genbank Database as Accession No. X68203.

The sequence encoding the first 12 amino acid residues was added to the expression construct by ligating an *SphI* digested PCR fragment amplified using reverse transcription-PCR of poly(A)$^+$ RNA isolated from the HEL cells using the oligonucleotides 5'-ACATGCATGC CACCATGCAG CGGGGCGCCG CGCTGTGCCT GCGACTGTGG CTCTGCCTGG GACTCCTGGA-3' (SEQ ID NO: 2)(forward primer, *SphI* site underlined, the translational start codon marked in bold follows an optimized Kozak consensus sequence Kozak, *Nucl. Acids Res. 15:* 8125-8148, 1987) and 5'-ACATGCATGC CCCGCCGGT CATCC-3' (SEQ ID NO: 3) (reverse primer, *SphI* site underlined) to the 5' end of the S2.5 fragment, thus replacing unique *SphI* fragment of the S2.5 plasmid. The resulting vector was digested with *EcoRI* and *ClaI* and ligated to a 138 bp PCR fragment amplified from the 0.6 kb *EcoRI* fragment (base pairs 3789 to 4416 in the Genbank X68203 sequence) which encodes the 3' end of Flt4s shown in Figure 1 of Pajusola *et al.*,

residues (MQRGAALCLRLW). Poly(A)$^+$ RNA was isolated from the human HEL erythroleukemia cells and double-stranded cDNA copy was synthesized using the Amersham cDNA Synthesis System Plus kit and a gene specific primer: 5'-TGTCCTCGCTGTCCTTGTCT-3' (SEQ ID NO: 1), which was located 195 bp downstream of the 5' end of clone S2.5. Double stranded cDNA was treated with T4 DNA polymerase to blunt the ends and cDNA was purified with Centricon 100 (Amicon Inc., Beverly, MA). Circularization of the blunt-ended cDNA was performed in a total volume of 150 microliters. The reaction mixture contained ligation buffer, 5% PEG-8000, 1 mM DTT and 8U of T4 DNA ligase (New England Biolabs). Ligation was carried out at 16°C for 16 hours. Fifteen microliters of this reaction mix were used in a standard 100 microliters PCR reaction containing 100 ng of specific primers including *SacI* and *PstI* restriction sites, present in this segment of the Flt4 cDNA, and 1 unit of Taq DNA polymerase (Perkin-Elmer Cetus). Two rounds of PCR were performed using 33 cycles (denaturation at 95°C for 1 minute, annealing at 55°C for 2 minutes and elongation at 72°C for 4 minutes). The PCR mixture was treated sequentially with the *SacI* and *PstI* restriction enzymes, and after purification with MagicPCR Preps (Promega), DNA fragments were subcloned into the pGEM3Zf(+) vector for sequencing. The sequence obtained corresponds to the 5' end of the Flt4s cDNA clone deposited in the Genbank Database as Accession No. X68203.

The sequence encoding the first 12 amino acid residues was added to the expression construct by ligating an *SphI* digested PCR fragment amplified using reverse transcription-PCR of poly(A)$^+$ RNA isolated from the HEL cells using the oligonucleotides 5'-ACATGCATGC CACCATGCAG CGGGGCGCCG CGCTGTGCCT GCGACTGTGG CTCTGCCTGG GACTCCTGGA-3' (SEQ ID NO: 2)(forward primer, *SphI* site underlined, the translational start codon marked in bold follows an optimized Kozak consensus sequence Kozak, *Nucl. Acids Res. 15:* 8125-8148, 1987) and 5'-ACATGCATGC CCCGCCGGT CATCC-3' (SEQ ID NO: 3) (reverse primer, *SphI* site underlined) to the 5' end of the S2.5 fragment, thus replacing unique *SphI* fragment of the S2.5 plasmid. The resulting vector was digested with *EcoRI* and *ClaI* and ligated to a 138 bp PCR fragment amplified from the 0.6 kb *EcoRI* fragment (base pairs 3789 to 4416 in the Genbank X68203 sequence) which encodes the 3' end of Flt4s shown in Figure 1 of Pajusola *et al.*,

Cancer Res. 52:5738-5743, 1992, using the oligonucleotides 5'-CGGAATTCCC CATGACCCCA AC-3' (SEQ ID NO: 4) (forward, *EcoRI* site underlined) and 5'-CCATCGATGG ATCCTACCTG AAGCCGCTTT CTT-3' (SEQ ID NO: 5) (reverse, *ClaI* site underlined). The coding domain was completed by ligation of the 1.2 kb *EcoRI* fragment (base pairs 2535-3789 of sequence X68203) into the above construct. The complete cDNA was subcloned as a *HindIII-ClaI(blunted)* fragment (this *ClaI* site was also included in the 3' primer used to construct the 3' end of the coding sequence) to the pLTRpoly expression vector reported in Mäkelä *et al.*, *Gene*, *118*: 293-294 (1992) (Genbank accession number X60280, SEQ ID NO: 37), incorporated by reference herein, using its *HindIII-Acc I(blunted)* restriction sites.

The long form of Flt4 was produced by replacing the 3'-end of the short form as follows: The 3' region of the Flt4l cDNA was PCR-amplified using a gene specific and a pGEM 3Z vector specific (SP6 promoter) oligonucleotide 5'-ATTTAGGTGACACTATA-3' (SEQ ID NO: 6) as reverse and forward primers, respectively, and an Flt4l cDNA clone containing a 495 bp *EcoRI* fragment extending downstream of the *EcoRI* site at nucleotide 3789 of the Genbank X68203 sequence (the sequence downstream of this *EcoRI* site is deposited as the Flt4 long form 3' sequence having Genbank accession number S66407 (SEQ ID NO: 38)). The gene specific oligonucleotide contained a *BamHI* restriction site located right after the end of the coding region. The sequence of that (reverse primer) oligonucleotide was 5'-CCATCGATGGATCCCGATGCTGCTTAGTAGCTGT-3' (SEQ ID NO: 7) (*BamHI* site is underlined). The PCR product was digested with *EcoRI* and *BamHI* and transferred in frame to LTRFlt4s vector fragment from which the coding sequences downstream of the EcoRI site at base pair 2535 (see sequence X68203) had been removed by *EcoRI-BamHI* digestion. Again, the coding domain was completed by ligation of the 1.2 kb *EcoRI* fragment (base pairs 2535-3789 of sequence X68203) back into the resulting construct.

EXAMPLE 2

Production and analysis of Flt4l transfected cells

NIH3T3 cells (60 % confluent) were co-transfected with 5 micrograms of the pLTRFlt4l construct and 0.25 micrograms of the pSV2neo

vector (ATCC) containing the neomycin phosphotransferase gene, using the DOTAP liposome-based transfection reagents (Boehringer Mannheim, Mannheim, Germany). One day after the transfection the cells were transferred into selection media containing 0.5 mg/ml geneticin (GIBCO, Grand Island, N.Y.). Colonies of geneticin-resistant cells were isolated and analyzed for expression of the Flt4 proteins. Cells were lysed in boiling lysis buffer containing 3.3% SDS (sodium dodecyl sulphate), 125 mM Tris, pH 6.8. Protein concentrations of the samples were measured by the BCA method (Pierce, Rockford, IL). About 50 micrograms protein of each lysate was analyzed for the presence of Flt4 by 6% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting using antisera against the carboxyl terminus of Flt4 and the ECL method (Amersham).

For production of anti-Flt4 antiserum the Flt4 cDNA fragment encoding the 40 carboxyterminal amino acid residues of the short form: NH2-PMTPTTYKG SVDNQTDSGM VLASEEFEQI ESRHRQESGFR-COOH (SEQ ID NO: 8) was cloned as a 657 bp *Eco*RI-fragment into the pGEX-1λT bacterial expression vector (Pharmacia) in frame with the glutathione-S-transferase coding region. The resultant GST-Flt4S fusion protein was produced in *E.coli* and purified by affinity chromatography using a glutathione-Sepharose 4B column. The purified protein was lyophilized, dissolved in phosphate buffered saline (PBS), mixed with Freund's adjuvant and used for immunization of rabbits at biweekly intervals using methods standard in the art (Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1988). Antisera were used after the fourth booster immunization for immunoprecipitation of Flt4 from the transfected cells and cell clones expressing Flt4 were used for ligand stimulation analysis.

EXAMPLE 3

Construction of a Flt4 EC baculovirus vector and expression and purification of its product

The construction of an Flt4 extracellular domain (EC) baculovirus vector is schematically depicted in Figure 3. The Flt4-encoding cDNA has been prepared in both a long form and a short form, each being incorporated in a vector under control of the Moloney murine leukemia virus LTR promoter. The nucleotide sequence of the short form of the Flt4 receptor

is available on the Genbank database as Accession No. X68203 and the specific 3' segment of the long form cDNA is available as Accession No. S66407.

The ends of a cDNA segment encoding Flt4 extracellular domain (EC) were modified as follows: The 3' end of Flt4 cDNA sequence (Genbank Accession Number X68203) which encodes the extracellular domain was amplified using primer 1116 5'-CTGGAGTCGACTTGGCGGACT-3' (SEQ ID NO: 9, *SalI* site underlined) and primer 1315 5'-CGCGGATCCCTAGTGATGGTGATGGTGATGTCTACCTTCGATCATGCT GCCCTTAT CCTC-3' (SEQ ID NO: 10, *BamHI* site underlined). The sequence complementary to that of primer 1315 continues after the Flt4 reading frame and encodes 6 histidine residues for binding to a Ni-NTA column (Qiagen, Hilden, Germany) followed by a stop codon, and an added Bam HI site. The amplified fragment was digested with *SalI* and *BamHI* and used to replace a unique *SalI-BamHI* fragment in the LTRFlt4 vector shown in Figure 3. The *SalI-BamHI* fragment that was replaced encodes the Flt4 transmembrane and cytoplasmic domains.

The 5' end without the Flt4 signal sequence encoding region was amplified by PCR using the primer 1335 5'-CCCAAGCTTGGATCCAAGTGGCTACTCCATGACC-3' (SEQ ID NO: 11) (the primer contains added *HindIII* (AAGCTT) and *BamHI* (GGATCC) restriction sites, which are underlined) and primer 1332 5'-GTTGCCTGTGATGTGCACCA-3' (SEQ ID NO: 12). The amplified fragment was digested with *HindIII* and *SphI* (the *HindIII* site (AAGCTT) is underlined in primer 1335 and the *SphI* site is within the amplified region of the Flt4l cDNA). The resultant *HindIII-SphI* fragment was used to replace a *HindIII-SphI* fragment in the modified LTRFlt4l vector described immediateiy above (the *HindIII* site is in the 5' junction of the Flt4 insert with the pLTRpoly portion of the vector, the *SphI* site is in Flt4 cDNA). The resultant Flt4EC insert was then ligated as a *BamHI* fragment into the *BamHI* site in the pVTBac plasmid as disclosed in Tessier *et al.*, *Gene 98*: 177-183 (1991), incorporated by reference herein. The orientation was confirmed to be correct by partial sequencing so that the open reading frame of the signal sequence-encoding portion of the vector continued in frame with the Flt4 sequence. That construct was transfected together with the baculovirus genomic DNA

into SF-9 cells by lipofection. Recombinant virus was purified, amplified and used for infection of High-Five cells (Invitrogen, San Diego, CA) using methods standard in the art. The Flt4 extracellular domain (Flt4EC) was purified from the culture medium of the infected High-Five cells using Ni-NTA affinity chromatography according to manufacturer's instructions (Qiagen) for binding and elution of the 6xHis tag encoded in the COOH-terminus of the recombinant Flt4 extracellular domain.

EXAMPLE 4

Isolation of Flt4 Ligand from Conditioned Media

An Flt4 ligand according to the invention was isolated from conditioned media from PC-3 prostatic adenocarcinoma cell line CRL1435 from the American Type Culture Collection and cultured as instructed by the supplier in Ham's F-12 Nutrient mixture (GIBCO) containing 7% fetal calf serum. In order to prepare the conditioned media, confluent PC-3 cells were cultured for 7 days in Ham's F-12 Nutrient mixture (GIBCO) in the absence of fetal bovine serum. Medium was then cleared by centrifugation at 10,000 g for 20 minutes. The medium was then screened to determine its ability to induce tyrosine phosphorylation of Flt4 by exposure to NIH3T3 cells which had been transfected with Flt4-encoding cDNA using the pLTRFlt4l vector. For receptor stimulation experiments, subconfluent NIH3T3 cells were starved overnight in serum-free DMEM medium (GIBCO) containing 0.2% BSA. The cells were stimulated with the conditioned media for 5 minutes, washed twice with cold PBS containing 100 micromolar vanadate and lysed in RIPA buffer (10 mM Tris pH 7.5, 50 mM NaCl, 0.5% sodium deoxycholate, 0.5% Nonidet P40 (BDH, Poole, England), 0.1% SDS, 0.1 U/ml Aprotinin (Boehringer Mannheim), 1 mM vanadate) for receptor immunoprecipitation analysis. The lysates were centrifuged for 20 minutes at 15,000 x g. The supernatants were incubated for 2 hours on ice with 3 microliters of the antiserum against the Flt4 C-terminus described in Example 2 and also in Pajusola, *et al. Oncogene 8*: 2931-2937 (1993), incorporated by reference herein.

After a 2 hour incubation in the presence of anti-Flt4 antiserum, protein A-Sepharose (Pharmacia) was added and incubation was continued for 45 minutes with rotation. The immunoprecipitates were washed three times

with the immunoprecipitation buffer and twice with 10 mM Tris, pH 7.5, before analysis in SDS-PAGE. Polypeptides were transferred to nitrocellulose and analyzed by Western blotting using Flt4- or phosphotyrosine-specific antisera and the ECL method (Amersham International, Buckinghamshire, England). Anti-phosphotyrosine monoclonal antibodies (anti-PTyr; PY20) were purchased from Transduction Laboratories (Lexington, Kentucky). In some cases, the filters were restained with a second antibody after stripping. The stripping of the filters was done for 30 minutes at 50°C in 100 mM 2-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl pH 6.7 with occasional agitation.

As shown in Figure 4, the PC-3 cell conditioned medium stimulated tyrosine phosphorylation of a 125 kD polypeptide when Flt4-expressing NIH3T3 cells were treated with the indicated preparations of media, lysed, and the lysates were immunoprecipitated with anti-Flt4 antiserum followed by SDS-PAGE, Western blotting, and staining using anti-PTyr antibodies. The resulting band was weakly phosphorylated upon stimulation with unconcentrated PC-3 conditioned medium (lane 2). The 125 kD band comigrated with the tyrosine phosphorylated, processed form of the mature Flt4 from pervanadate-treated cells (compare lanes 2 and 7 of Fig. 4, see also Figure 5A). Comigration was confirmed upon restaining with anti-Flt4 antibodies as is also shown in Figure 5A (panel on the right). In order to show that the 125 kD polypeptide is not a non-specific component of the conditioned medium reactive with anti-phosphotyrosine antibodies, 15 microliters of conditioned medium was separated by SDS-PAGE, blotted on nitrocellulose and the blot was stained with anti-PTyr antibodies. No signal was obtained (Fig. 5B). Also, unconditioned medium failed to stimulate Flt4 phosphorylation, as shown in Figure 4, lane 1.

Figure 5C shows a comparison of the effects of PC-3 CM stimulation (+) on untransfected (lanes 4 and 5), FGFR-4-transfected (lanes 8 and 9) and Flt4-transfected NIH 3T3 cells (lanes 1-3, 6 and 7). These results indicate that neither untransfected NIH 3T3 cells nor NIH 3T3 cells transfected with FGF receptor 4 showed tyrosine phosphorylation of p120 upon stimulation with the conditioned medium from PC-3 cells. Analysis of stimulation by PC-3 CM pretreated with Heparin Sepharose CL-6B

(Pharmacia) for 2 hours at room temperature (lane 3) shows that the Flt4 ligand does not bind to heparin.

As shown in Figure 4, lane 3, stimulating activity was considerably increased when the PC-3 conditioned medium was concentrated four-fold using a Centricon-10 concentrator (Amicon). Figure 4, lane 4, shows that pretreatment of the concentrated PC-3 conditioned medium with 50 microliters of the Flt4 extracellular domain coupled to CNBr-activated sepharose CL-4B (Pharmacia; about 1mg of Flt4EC domain/ml sepharose resin) completely abolished Flt4 tyrosine phosphorylation. Similar pretreatment of the conditioned medium with unsubstituted sepharose CL-4B did not affect stimulatory activity, as shown in Figure 4, lane 5. Also, the flow through obtained after concentration, which contained proteins of less than 10,000 molecular weight, did not stimulate Flt4 phosphorylation, as shown in Figure 4, lane 6.

In another experiment, a comparison of Flt4 autophosphorylation in transformed NIH 3T3 cells expressing LTRFlt4l was conducted, using unconditioned medium, medium from PC-3 cells expressing the Flt4 ligand, or unconditioned medium containing either 50 ng/ml of VEGF165 or 50 ng/ml of PlGF-1. The cells were lysed, immunoprecipitated using anti-Flt4 antiserum and analyzed by Western blotting using anti-phosphotyrosine antibodies. As shown in Fig. 8, only the PC-3 conditioned medium expressing the Flt4 ligand (lane Flt-4L) stimulated Flt4 autophosphorylation.

The foregoing data show that PC-3 cells produce a ligand which binds to the extracellular domain of Flt4 and activates this receptor.

EXAMPLE 5

Purification of the Flt4 Ligand

The ligand expressed by PC-3 cells as characterized in Example 3 was purified and isolated using a recombinantly-produced Flt4 extracellular domain (Flt4EC) in affinity chromatography.

Two harvests of serum-free conditioned medium, comprising a total of 8 L, were collected from 500 confluent 15 cm diameter culture dishes containing confluent layers of PC-3 cells. The conditioned medium was clarified by centrifugation at 10,000 x g and concentrated 80-fold using an

Ultrasette Tangential Flow Device (Filtron, Northborough, MA) with a 10 kD cutoff Omega Ultrafiltration membrane according to the manufacturer's instructions. Recombinant Flt4 extracellular domain was expressed in a recombinant baculovirus cell system and purified by affinity chromatography on Ni-agarose (Ni-NTA affinity column obtained from Qiagen). The purified exracellular domain was coupled to CNBr-activated Sepharose CL-4B at a concentration of 5 mg/ml and used as an affinity matrix for ligand affinity chromatography.

Concentrated conditioned medium was incubated with 2 ml of the recombinant Flt4 extracellular domain-Sepharose affinity matrix in a rolling tube at room temperature for 3 hours. All subsequent purification steps were at +4 C. The affinity matrix was then transferred to a column (Pharmacia) with an inner diameter of 15 mm and washed successively with 100 ml of PBS and 50 ml of 10 mM Na-phosphate buffer (pH 6.8). Bound material was eluted step-wise with 100 mM glycine-HCl, successive 6 ml elutions having pHs of 4.0, 2.4, and 1.9. Several 2 ml fractions of the eluate were collected in tubes containing 0.5 ml 1 M Na-phosphate (pH 8.0). Fractions were mixed immediately and dialysed in 1 mM Tris-HCl (pH 7.5). Aliquots of 75 ul each were analyzed for their ability to stimulate tyrosine phosphorylation of Flt4. The ultrafiltrate, 100 ul aliquots of the concentrated conditioned medium before and after ligand affinity chromatography, as well as 15-fold concentrated fractions of material released from the Flt4 extracellular domain-Sepharose matrix during the washings were also analyzed for their ability to stimulate Flt4 tyrosine phosphorylation.

As shown in Figure 6, lane 3, the concentrated conditioned medium induced prominent tyrosine phosphorylation of Flt4 in transfected NIH3T3 cells overexpressing Flt4. This activity was not observed in conditioned medium taken after medium was exposed to the Flt4 Sepharose affinity matrix described above (lane 4). The specifically-bound Flt4-stimulating material was retained on the affinity matrix upon washes in PBS, 10 mM Na-phosphate buffer (pH 6.8), and at pH 4.0 (lanes 5-7, respectively), and it was eluted in the first two 2 ml aliquots at pH 2.4 (lanes 8 and 9). A further decrease of the pH of the elution buffer did not cause release of additional Flt4-stimulating material (lane 11). Lane 1 depicts a control wherein Flt4-expressing cells were treated with nonconditioned medium; lane 2

depicts results wherein Flt4-expressing cells were treated with the ultrafiltrate fraction of conditioned medium containing polypeptides of less than 10 kD molecular weight.

Small aliquots of the chromatographic fractions were concentrated in a SpeedVac concentrator (Savant, Farmingdale, N.Y.) and subjected to SDS-PAGE under reducing conditions with subsequent silver staining of the gel. As shown in Figure 7, the major polypeptide, having a molecular weight of approximately 23 kD (reducing conditions), was detected in the fractions containing Flt4 stimulating activity (corresponding to lanes 8 and 9 in Figure 6). That polypeptide was not found in the other chromatographic fractions. On the other hand, all other components detected in the two active fractions were also distributed in the starting material and in small amounts in the other washing and elution steps after their concentration. Similar results were obtained in three independent affinity purifications, indicating that the 23 kD polypeptide specifically binds to Flt4 and induces its tyrosine phosphorylation.

Fractions containing the 23 kD polypeptide were combined, dried in a SpeedVac concentrator and subjected to SDS-PAGE in a 12.5 % gel. The proteins from the gel were then electroblotted to Immobilon-P (PVDF) transfer membrane (Millipore, Malborough, MA) and visualized by staining of the blot with Coomassie blue R-250. The region containing only the stained 23 kD band was cut from the blot and was subjected to N-terminal amino acid sequence analysis in a Prosite Protein Sequencing System (Applied Biosystems, Foster City, CA). The data were analyzed using a 610A Data Analysis System (Applied Biosystems). Analysis revealed a single N-terminal sequence of NH_2-XEETIKFAAAHYNTEILK-COOH (SEQ ID NO: 13).

EXAMPLE 6

Construction of PC-3 cell cDNA library in a eukaryotic expression vector.

Poly(A)$^+$ RNA was isolated from five 15 cm diameter confluent dishes of PC-3 cells by a single step method using oligo(dT) (Type III, Collaborative Research) cellulose affinity chromatography (Sambrook et al., Molecular Cloning, A Laboratory Manual; Cold Spring Harbor Laboratory Press, 1989). The yield was 70 micrograms. Six micrograms of the Poly(A)$^+$

RNA was used to prepare an oligo(dT)-primed cDNA library in the mammalian expression vector pcDNA I and the Librarian kit of Invitrogen according to the instructions included in the kit. The library was estimated to contain about 10^6 independent recombinants with an average insert size of approximately 1.8 kb.

EXAMPLE 7

Amplification of the unique nucleotide sequence encoding the Flt4 ligand

Degenerate oligonucleotides were designed based on the N-terminal amino acid sequence of the isolated Flt4 ligand and were used as primers in a polymerase chain reaction (PCR) to amplify cDNA encoding the Flt4 ligand from a PC-3 cell library. The overall strategy is schematically depicted in Fig. 9A, where the different primers have been marked with arrows.

The PCR was carried out using 1 microgram of DNA from the amplified PC-3 cDNA library and a mixture of sense-strand primers comprising 5'-GCAGARGARACNATHAA-3' (SEQ ID NO: 14)(wherein R is A or G, N is A,G,C or T and H is A, C or T), encoding amino acid residues 2-6 (EETIK, SEQ ID NO: 15) and antisense-strand primers 5'-GCAYTTNARDATYTCNGT-3' (SEQ ID NO: 16) (wherein Y is C or T and D is A, G or T), corresponding to amino acid residues 14-18 (TEILK, SEQ ID NO: 17). Three extra nucleotides (GCA) were added to the 5'-terminus of each primer to increase annealing stability. Two successive PCR runs were carried out using 1 U per reaction of DynaZyme (F-500L, Finnzymes), a thermostable DNA polymerase, in a buffer supplied by the manufacturer (10 mM Tris-HCl, pH 8.8 at 25 °C, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton-X100), at an extension temperature of 72 °C. The first PCR run was carried out for 43 cycles. The first three cycles were run at an annealing temperature of 33 °C for 2 minutes, and the remaining cycles were run at 42 °C for 1 minute.

The region of the gel containing a weak band of the expected size (57 bp) was cut out from the gel and eluted. The eluted material was reamplified for 30 cycles using the same primer pairs described above at 42°C for 1 minute. The amplified fragment was cloned into a pCR II vector

(Invitrogen) using the TA cloning kit (Invitrogen) and sequenced using the radioactive dideoxynucleotide sequencing method of Sanger. Six clones were analysed and all contained the sequence encoding the expected peptide (amino acid residues 2-18 of the Flt4 ligand precursor). Nucleotide sequence spanning the region from the third nucleotide of codon 6 to the third nucleotide of codon 13 (the extension region) was identical in all six clones: 5'-ATTCGCTGCAGCACACTACAAC-3' (SEQ ID NO: 18) and thus was considered to represent an amplified product from the unique sequence encoding part of the amino terminus of the Flt4 ligand.

EXAMPLE 8

Amplification of the 5'-end of the cDNA encoding the Flt4 ligand

Based on the unique nucleotide sequence encoding the N-terminus of the isolated Flt4 ligand, two pairs of nested primers were designed to amplify, in two subsequent PCR-reactions, the complete 5'-end of the corresponding cDNAs from one microgram of DNA from the above-described PC-3 cDNA library. First, amplification was performed with primer 5'-TCNGTGTTGTAGTGTGCTG-3' (SEQ ID NO: 19), which is the antisense-strand primer corresponding to amino acid residues 9-15 (AAHYNTE, SEQ ID NO: 20), and sense-strand primer 5'-TAATACGACTCACTATAGGG-3' (SEQ ID NO: 21), corresponding to the T7 RNA promoter of the pcDNAI vector used for construction of the library. "Touchdown" PCR was used as disclosed in Don, *et al.*, *Nucl. Acids Res.*, *19*: 4008 (1991), incorporated by reference herein. The annealing temperature of the two first cycles was 62 °C and subsequently the annealing temperature was decreased in every other cycle by 1 °C until a final temperature of 53 °C was reached, at which temperature 16 additional cycles were conducted. Annealing time was . minute and extension at each cycle was conducted at 72 °C for 1 minute. Multiple amplified DNA fragments were obtained in the first reaction. The products of the first amplification (1 ul of a 1:100 dilution in water) were used in the second amplification reaction employing the nested primers 5'-GTTGTAGTGTGCTGCAGCGAATTT-3' (SEQ ID NO: 22), an antisense-strand primer corresponding to amino acid residues 6-13 (KFAAAHYN, SEQ ID NO: 23) of the Flt4 ligand, and 5'-TCACTATAGGGAGACCCAAGC-3'

(SEQ ID NO: 24), a sense-strand primer corresponding to nucleotides 2179-2199 of the pcDNAI vector. The sequences of these sense and antisense primers overlapped with the 3' ends of the corresponding primers used in the first PCR. "Touchdown" PCR was carried out by decreasing the annealing temperature from 72 °C to 66 °C and continuing with 18 additional cycles at 66 °C. The annealing time was 1 minute and extension at each cycle was carried out at 72 °C for 2 minutes. One major product of about 220 bp and three minor products of about 270 bp, 150 bp, and 100 bp were obtained.

The amplified fragment of approximately 220 bp was cut out from the agarose gel, cloned into a pCRII vector using the TA cloning kit (Invitrogen) and sequenced. Three recombinant clones were analysed and they contained the sequence 5'-
TCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGT
AACGGCCGCCAGTGTGGTGGAATTCGACGAACTCATGACTGTACTCT
ACCCAGAATATTGGAAAATGTACAAGTGTCAGCTAAGGCAAGGAGGC
TGGCAACATAACAGAGAACAGGCCAACCTCAACTCAAGGACAGAAG
AGACTATAAAATTCGCTGCAGCACACTACAAC- 3' (SEQ ID NO: 25).
The beginning of the sequence represents the pcDNAI vector and the underlined sequence represents the amplified product of the 5'-end of the insert.

EXAMPLE 9

Amplification of the 3'-end of cDNA encoding the Flt4 ligand

Based upon the amplified 5'-sequence of the clones encoding the amino terminus of the 23 kD Flt4 ligand, two pairs of non-overlapping nested primers were designed to amplify the 3'-portion of the Flt-4-ligand-encoding cDNA clones. The sense-strand primer 5'-ACAGAGAACAGGCCAACC-3' (SEQ ID NO: 26) and antisense-strand primer 5'-TCTAGCATTTAGGTGACAC-3' (SEQ ID NO: 27) corresponding to nucleotides 2311-2329 of the pcDNAI vector were used in a first "touchdown" PCR. The annealing temperature of the reaction was decreased 1°C every two cycles from 72°C to 52°C, at which temperature 15 additional cycles were carried out. The annealing time was 1 minute and extension at each cycle was carried out at 72°C for 3 minutes. DNA fragments of several sizes were obtained in the first amplification. Those products were diluted 1:200 in water

and reamplified in PCR using the second pair of primers: 5'-AAGAGACTATAAAATTCGCTGCAGC-3' (SEQ ID NO: 28) and 5'-CCCTCTAGATGCATGCTCGA-3' (SEQ ID NO: 29) (antisense-strand primer corresponding to nucleotides 2279-2298 of the pcDNAI vector). Two DNA fragments were obtained, having sizes of 1350 bp and 570 bp. Those fragments were cloned into a pCRII vector and the inserts of the clones were sequenced. Both of these fragments were found to contain sequences encoding an amino acid sequence homologous to the VEGF sequence.

EXAMPLE 10

Screening the PC-3 cell cDNA library using the 5' PCR fragment of Flt4 ligand cDNA

A 219 bp 5'-terminal fragment of Flt4 ligand cDNA was amplified by PCR using the 5' PCR fragment described above and primers 5'-GTTGTAGTGTGCTGCAGCGAATTT-3' (antisense-strand primer, SEQ ID NO: 30) and 5'-TCACTATAGGGAGACCCAAGC-3' (SEQ ID NO: 31) (sense-primer corresponding to nucleotides 2179-2199 of the pcDNAI vector). The amplified product was subjected to digestion with *EcoRI* (Boehringer Mannheim) to remove the portion of the DNA sequence amplified from the pcDNAI vector and the resulting 153 bp fragment encoding the 5' end of the Flt4 ligand was labeled with [^{32}P]-dCTP using the Klenow fragment of *E.coli* DNA polymerase I (Boehringer Mannheim). That fragment was used as a probe for hybridization screening of the amplified PC-3 cell cDNA library.

Filter replicas of the library were hybridized with the radioactively labeled probe at 42 °C for 20 hours in a solution containing 50% formamide, 5x SSPE, 5x Denhardt's solution, 0.1% SDS and 0.1 mg/ml denatured salmon sperm DNA. Filters were washed twice in 1x SSC, 0.1% SDS for 30 minutes at room temperature, then twice for 30 minutes at 65 °C and exposed overnight.

On the basis of autoradiography, 10 positive recombinant bacterial colonies hybridizing with the probe were chosen from the library. Plasmid DNA was purified from these colonies and analysed by *EcoRI* and *NotI* digestion and agarose gel electrophoresis followed by ethidium bromide staining. The ten plasmid clones were divided into three groups on the basis of the presence of insert sizes of approximately 1.7, 1.9 and 2.1 kb.

respectively. Inserts of plasmids from each group were sequenced using the T7 oligonucleotide as a primer and walking primers for subsequent sequencing reactions.

Sequence analysis showed that all clones contain the open reading frame encoding the NH2-terminal sequence of the 23 kD Flt4 ligand. Dideoxy sequencing was continued using walking primers in the downstream direction. A complete cDNA sequence and deduced amino acid sequence from a 2.1 kb clone is set forth in Fig. 9B (SEQ ID NOs: 32 and 33, respectively). A putative cleavage site of a "prepro" leader sequence is indicated in Fig. 9B with a shaded triangle. When compared with sequences in the GenBank Database, the predicted protein product of this reading frame was found to be homologous with the predicted amino acid sequences of the PDGF/VEGF family of growth factors, as shown in Figure 10.

Plasmid pFLT4-L, containing the 2.1 kb cDNA clone in pcDNAI vector, has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 as accession number 97231.

EXAMPLE 11

Stimulation of Flt4 autophosphorylation by the protein product of the Flt4 ligand vector

The 2.1 kb cDNA insert of plasmid Flt4-L, which contains the open reading frame encoding the sequence shown in Fig. 9B (SEQ ID NO: 32), was cut out from the pcDNAI vector using *HindIII* and *NotI* restriction enzymes, isolated from a preparative agarose gel, and ligated to the corresponding sites in the pREP7 expression vector (Invitrogen). The pREP7 vector containing the above cloned pFLT4-L insert was transfected into 293-EBNA cells (Invitrogen) using the calcium phosphate transfection method (Sambrook et al., Molecular Cloning, A Laboratory Manual; Cold Spring Harbor Laboratory Press, 1989). About 48 hours after transfection the medium of the transfected cells was changed to DMEM medium lacking fetal calf serum and incubated for 36 h. The thus conditioned medium was then collected, centrifuged at 5000 x g for 20 minutes, the supernatant was concentrated 5-fold using Centriprep 10 (Amicon) and used to stimulate NIH3T3 cells expressing LTRFlt4l, as in Example 4. The cells were lysed,

immunoprecipitated using anti-Flt4 antiserum and analysed by Western blotting using anti-phosphotyrosine antibodies.

As can be seen from Fig. 11, lanes 1 and 3, the conditioned medium from two different dishes of the transfected cells stimulated Flt4 autophosphorylation in comparison with the medium from mock-transfected cells, which gave only background levels of phosphorylation of the Flt4 receptor (lane 2). When the concentrated conditioned medium was pre-absorbed with 20 microliters of a slur of Flt4EC domain coupled to Sepharose (see example 4), no phosphorylation was obtained (lane 4), showing that the activity responsible for Flt4 autophosphorylation was indeed the Flt4 ligand. Thus, these results demonstrate that an expression vector having an approximately 2.1 kb insert and containing an open reading frame as shown in Fig. 9B (and SEQ ID NO:32) is expressed as a biologically active Flt4 ligand in transfected cells. The sequence encoded by that open reading frame is shown in SEQ ID NO: 33.

The deduced molecular weight of a polypeptide consisting of the complete amino acid sequence in Fig. 9B (SEQ ID NO: 33, residues -102 to 317) is 46,883. The deduced molecular weight of a polypeptide consisting of amino acid residues 1 to 317 of SEQ ID NO: 33 is 35,724. The Flt4 ligand purified from PC-3 cultures had an observed molecular weight of about 23 kD as assessed by SDS-PAGE under reducing conditions. Thus, it appears that the Flt4 ligand mRNA is translated into a precursor polypeptide, from which the mature ligand is derived by proteolytic cleavage. Also, the Flt4 ligand may be glycosylated at three putative N-linked glycosylation sites conforming to the consensus which can be identified in the deduced Flt4 ligand amino acid sequence (N-residues underlined in Fig. 10).

The carboxyl terminal amino acid sequences, which increase the predicted molecular weight of the Flt4 ligand subunit in comparison with other ligands of this family, show a pattern of spacing of cysteine residues reminiscent of the Balbiani ring protein 3 (BRP3) sequence (Dignam and Case, *Gene*, *88*:133-140, 1990), as depicted schematically in Fig. 9A. Such a sequence may encode an independently folded domain present in a Flt4 ligand precursor and it may be involved, for example, in the regulation of secretion, solubility, stability, cell surface localization or activity of the Flt4 ligand.

Interestingly, at least one cysteine motif of the BRP3 type is also found in the VEGF carboxy terminal amino acid sequences.

Thus, the Flt4 ligand mRNA appears first to be translated into a precursor from the mRNA corresponding to the cDNA insert of plasmid FLT4-L, from which the mature ligand is derived by proteolytic cleavage. To define the mature Flt4 ligand polypeptide one first expresses the cDNA clone (which is deposited in the pcDNAI expression vector) in cells, such as COS cells. One uses antibodies generated against encoded peptides, fragments thereof, or bacterial Flt4 fusion proteins, such as a GST-fusion protein, to raise antibodies against the VEGF-homologous domain and the amino- and carboxyl-terminal propeptides of Flt4 ligand. One then follows the biosynthesis and processing of the Flt4 ligand in the transfected cells by pulse-chase analysis using radioactive cysteine for labelling of the cells, immunoprecipitation and gel electrophoresis. Using antibodies against the three domains of the product encoded by the cDNA insert of plasmid FLT4-L, material for radioactive or nonradioactive amino-terminal sequence analysis is isolated. The determination of the amino-terminal sequence of the mature VEGF-C polypeptide allows for identification of the amino-terminal proteolytic processing site. The determinmation of the amino-terminal sequence of the carboxyl-terminal propeptide will give the carboxyl-terminal processing site. This is confirmed by site-directed mutagenesis of the amino acid residues adjacent to the cleavage sites, which would prevent the cleavage.

On the other hand, the Flt4 ligand is characterized by progressive 3' deletions in the 3' coding sequences of the Flt4 ligand precursor clone, introducing a stop codon resulting in carboxy-terminal truncations of its protein product. The activities of such truncated forms are assayed by, for example, studying Flt4 autophosphorylation induced by the truncated proteins when applied to cultures of cells, such as NIH3T3 cells expressing LTRFlt4. By extrapolation from studies of the structure of the related platelet derived growth factor (PDGF, reference Heldin *et al.*, *Growth Factors 8*:245-252 (1993)) one determines that the region critical for receptor activation by the Flt4 ligand is contained within its first approximately 180 amino acid residues of the secreted VEGF-C protein lacking the putative 102 amino acid prepro leader, and apparently within the first approximately 120 amino acid residues.

On the other hand, the difference between the molecular weights observed for the purified ligand and deduced from the open reading frame of the Flt4 precursor clone may be due to the fact that the soluble ligand was produced from an alternatively spliced mRNA which would also be present in the PC-3 cells, from which the isolated ligand was derived. To isolate such alternative cDNA clones one uses cDNA fragments of the deposited clone and PCR primers made according to the sequence provided as well as techniques standard in the art to isolate or amplify alternative cDNAs from the PC-3 cell cDNA library. One may also amplify using reverse transcription (RT)-PCR directly from the PC-3 mRNA using the primers provided in the sequence of the cDNA insert of plasmid FLT4-L. Alternative cDNA sequences are determined from the resulting cDNA clones. One can also isolate genomic clones corresponding to the Flt4 ligand mRNA transcript from a human genomic DNA library using methods standard in the art and to sequence such clones or their subcloned fragments to reveal the corresponding exons. Alternative exons can then be identified by a number of methods standard in the art, such as heteroduplex analysis of cDNA and genomic DNA, which are subsequently characterized.

EXAMPLE 12

Expression of the Gene Encoding VEGF-C in Human Tumor Cell Lines

Expression of transcripts corresponding to the Flt4 ligand (VEGF-C) was analysed by hybridization of Northern blots containing isolated poly(A)$^+$ RNA from HT-1080 and PC-3 human tumor cell lines. The probe was the radioactively labelled insert of the 2.1 kb cDNA clone (specific activity 10^8-10^9 cpm/mg of DNA). The blot was hybridized overnight at 42 C using 50% formamide, 5 x SSPE buffer, 2% SDS, 10 x Denhardt's solution, 100 mg/ml salmon sperm DNA and 1 x 10^6 cpm of the labelled probe/ml. The blot was washed at room temperature for 2 x 30 minutes in 2 x SSC containing 0.05% SDS, and then for 2 x 20 min at 52 C in 0.1 x SSC containing 0.1% SDS. The blot was then exposed at -70 C for three days using intensifying screens and Kodak XAR film. Both cell lines expressed an Flt4 ligand mRNA of about 2.4 kb, as well as VEGF and VEGF-B mRNA's (Fig. 12).

EXAMPLE 13

VEGF-C Chains Are Proteolytically Processed after Biosynthesis and Disulfide Linked

The predicted molecular mass of the secreted polypeptide, as deduced from the VEGF-C ORF, is 46,883 kD, suggesting that VEGF-C mRNA may be first translated into a precursor, from which the mature ligand of 23 kD is derived by proteolytic cleavage.

To study this, metabolical labelling of 293 EBNA cells transfected with the VEGF-C construct was carried out by addition of 100 mCi/ml of Pro-mix™ L-[^{35}S] *in vitro* cell labelling mix (Amersham) to the culture medium devoid of cysteine and methionine. After two hours, the cell layers were washed twice with PBS and the medium was then replaced with DMEM-0.2% BSA. After 1, 3, 6, 12 and 24 hours of subsequent incubation, the culture medium was collected, clarified by centrifugation, and concentrated, and VEGF-C was bound to 30 microliters of a slur of Flt4EC-Sepharose overnight at +4 C, followed by three washes in PBS, two washes in 20 mM Tris-HCl (pH 7.5), alkylation, SDS-PAGE and autoradiography.

These experiments demonstrated that a putative precursor polypeptide of 32 kD apparent molecular mass was bound to the Flt4EC affinity matrix from the CM of metabolically labelled cells transfected with a VEGF-C expression vector (Fig. 13A). Increased amounts of a 23 kD receptor binding polypeptide accumulated in the culture medium during a subsequent chase period of 3 h, but not thereafter (lanes 2-4 and data not shown), suggesting that the 23 kD form is produced by proteolytic processing, which is cell-associated and incomplete, at least in the transiently transfected cells. The arrows in Fig. 13A indicate the 32 kDa and 23 kDa polypeptides of secreted VEGF-C.

In a related experiment, VEGF-C isolated using Flt4EC-Sepharose after a 4 h continuous metabolic labelling was analyzed by polyacrylamide gel electrophoresis in nonreducing conditions (Fig. 13B). Higher molecular mass forms were observed under nonreducing conditions, suggesting that the VEGF-C polypeptides can form disulfide-linked dimers and/or multimers (arrows in Fig. 13B).

EXAMPLE 14

Stimulation Of VEGFR-2 Autophosphorylation By VEGF-C

Conditioned medium (CM) from 293 EBNA cells, transfected with the VEGF-C vector also was used to stimulate porcine aortic endothelial (PAE) cells expressing VEGFR-2. Pajusola et al., *Oncogene*, *9*:3545-55 (1994); Waltenberger et al., *J. Biol. Chem.*, *269*:26988-95 (1994). The cells were lysed and immunoprecipitated using VEGFR-2 - specific antiserum (Waltenberger et al., 1994).

PAE-KDR cells (Waltenberger et al., 1994) were grown in Ham's F12 medium-10% fetal calf serum (FCS). Confluent NIH3T3-Flt4 cells or PAE-KDR cells were starved overnight in DMEM or Ham's F12 media, respectively, supplemented with 0.2% bovine serum albumin (BSA), and then incubated for 5 min. with the analyzed media. Recombinant human VEGF (R&D Systems) and PDGF-BB were used as a control stimulating agents. The cells were washed twice with ice-cold tris-buffered saline (TBS) containing 100 mM sodium orthovanadate and lysed in RIPA buffer containing 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.1 U/ml aprotinin and 1 mM sodium orthovanadate. The lysates were sonicated, clarified by centrifugation at 16,000 g for 20 min. and incubated for 3-6 h on ice with 3-5 microliters of antisera specific for Flt4 (Pajusola et al., 1993), VEGFR-2 or PDGFR-β (Claesson-Welsh et al., *J. Biol. Chem.*, *264*:1742-47 (1989); Waltenberger et al., 1994). Immunoprecipitates were bound to protein A-Sepharose, washed three times with RIPA buffer containing 1mM PMSF, 1mM sodium orthovanadate. twice with 10 mM Tris-HCl (pH 7.4) and subjected to SDS-PAGE in a 7% gel. Polypeptides were transferred to nitrocellulose by Western blotting and analysed using PY20 phosphotyrosine-specific monoclonal antibodies (Transduction Laboratories) or receptor-specific antiserum and ECL method (Amersham).

The results of the experiment are presented in Figs. 14A and 14B. As shown in Fig. 14A, PAE cells expressing VEGFR-2 were stimulated with 10- or 2-fold concentrated medium from mock-transfected 293-EBNA cells (lanes 1 and 2), or with 2-, 5- or 10-fold concentrated medium from 293 EBNA cell cultures expressing the recombinant VEGF-C (lanes 3-6). VEGFR-2 was immunoprecipitated with specific antibodies and analysed by SDS-PAGE and Western blotting using phosphotyrosine antibodies. For

comparison, the stimulations were carried out with non-conditioned medium containing 50 ng/ml of purified recombinant VEGF (lanes 7 and 8). Lanes 6 and 7 show stimulation with VEGF-C- or VEGF- containing media pretreated with Flt4EC. As depicted in Fig. 14B, PDGFR-β-expressing NIH3T3 cells were stimulated with non-conditioned medium (lane 1), 5-fold concentrated CM from mock-transfected (lane 2) or VEGF-C - transfected (lanes 3 and 4) cells, or with non-conditioned medium containing 50 ng/ml of recombinant human PDGF-BB (lane 5). Medium containing VEGF-C was also pretreated with recombinant Flt4EC (lane 4). PDGFR-β was immunoprecipitated with specific antibodies and analysed by SDS-PAGE and Western blotting using phosphotyrosine antibodies with subsequent stripping and reprobing of the membrane with antibodies specific for PDGFR-β.

Referring again to Fig. 14A, a basal level of tyrosine phosphorylation of VEGFR-2 was detected in cells stimulated by CM from the mock-transfected cells. A further concentration of this medium resulted in only a slight enhancement of VECFR-2 phosphorylation (lanes 1 and 2). CM containing recombinant VEGF-C stimulated tyrosine autophosphorylation of VEGFR-2 and the intensity of the autophosphorylated polypeptide band was increased upon concentration of the VEGF-C CM (lanes 3-5). Furthermore, the stimulating effect was abolished after pretreatment of the medium with the Flt4EC affinity matrix (compare lanes 1, 5 and 6). The maximal effect of VEGF-C in this assay was comparable to the effect of recombinant VEGF added to unconditioned medium at concentration of 50 ng/ml (lane 8). Pretreatment of the medium containing VEGF with Flt4EC did not abolish its stimulating effect on VEGFR-2 (compare lanes 7 and 8). These results suggest that the VEGF-C expression vector encodes a ligand not only for Flt4 (VEGFR-3), but also for VEGFR-2.

In order to further confirm that the stimulating effect of VEGF-C on tyrosine phosphorylation of VEGFR-3 and VEGFR-2 was receptor-specific, we analysed the effect of VEGF-C on tyrosine phosphorylation of PDGF receptor β (PDGFR-β) which is abundantly expressed on fibroblastic cells. As can be seen from Fig. 14B, a weak tyrosine phosphorylation of PDGFR-β was detected upon stimulation of Flt4-expressing NIH3T3 cells with CM from the mock-transfected cells (compare lanes 1 and 2). A similar low level of PDGFR-β phosphorylation

was observed when the cells were incubated with CM from the VEGF-C transfected cells, with or without prior treatment with Flt4EC (lanes 3 and 4). In contrast, the addition of 50 ng/ml of PDGF-BB induced a prominent tyrosine autophosphorylation of PDGFR-β (lane 5).

EXAMPLE 15

VEGF-C Stimulates Endothelial Cell Migration In Collagen Gel

CM from cell cultures transfected with the VEGF-C expression vector was placed in a well made in collagen gel and used to stimulate the migration of bovine capillary endothelial (BCE) cells in the three-dimensional collagen gel as follows.

BCE cells (Folkman et al., *Proc. Nat'l Acad. Sci. USA, 76*:5217-5221 (1979) were cultured as described in Pertovaara et al., *J. Biol. Chem., 269*:6271-74 (1994). The collagen gels were prepared by mixing type I collagen stock solution (5 mg/ml in 1 mM HCl) with an equal volume of 2x MEM and 2 volumes of MEM containing 10% newborn calf serum to give a final collagen concentration of 1.25 mg/ml. The tissue culture plates (5 cm diameter) were coated with about 1 mm thick layer of the solution, which was allowed to polymerize at 37°C. BCE cells were seeded on top of this layer. For the migration assays, the cells were allowed to attach inside a plastic ring (1 cm diameter) placed on top of the first collagen layer. After 30 min., the ring was removed and unattached cells were rinsed away. A second layer of collagen and a layer of growth medium (5% newborn calf serum (NCS)), solidified by 0.75% low melting point agar (FMC BioProducts, Rockland, ME), were added. A well (3 mm diameter) was punched through all the layers on both sides of the cell spot at a distance of 4 mm, and the sample or control media were pipetted daily into the wells. Photomicrographs of the cells migrating out from the spot edge were taken after six days through an Olympus CK 2 inverted microscope equipped with phase-contrast optics. The migrating cells were counted after nuclear staining with the fluorescent dye bisbenzimide (1 mg/ml, Hoechst 33258, Sigma).

Fig. 15A depicts a comparison of the number of cells migrating at different distances from the original area of attachment towards wells containing media conditioned by the non-transfected (control) or transfected

(mock; VEGF-C; VEGF) cells, 6 days after addition of the media. The number of cells migrating out from the original ring of attachment was counted in five adjacent 0.5 mm x 0.5 mm squares using a microscope ocular lens grid and 10x magnification. Cells migrating further than 0.5 mm were counted in a similar way by moving the grid in 0.5 mm steps. The experiments were carried out twice with similar results, and medium values from the one of the experiments are presented with standard error bars. The photographs in Fig. 15B depict phase-contrast microscopy and fluorescent microscopy of the nuclear staining of BCE cells migrating towards the wells containing media conditioned by the mock-transfected cells or by VEGF-C - transfected cells. The areas shown is approximately 1mm x 1.5mm, and arrows indicate the borders of the original ring of attachment.

After 6 days of treatment, the cultures were stained and cells at different distances outside of the original ring of attachment were counted using fluorescent nuclear staining and detection with a fluorescence microscope equipped with a grid. A comparison of the numbers of migrating cells in successive 0.5 mm x 0.5 mm areas is shown in Fig 15A. As can be seen from the columns, VEGF-C-containing CM stimulated cell migration more than medium conditioned by the non-transfected or mock-transfected cells but less than medium from cells transfected with a VEGF expression vector. An example of typical phase contrast and fluorescent microscopic fields of cultures stimulated with medium from mock-transfected or VEGF-C transfected cells is shown in Fig. 15B. Daily addition of 1 ng of FGF2 into the wells resulted in the migration of approximately twice the number of cells when compared to the stimulation by CM from VEGF-transfected cells.

EXAMPLE 16

VEGF-C Is Expressed In Multiple Tissues

Northern blots containing 2 micrograms of isolated poly(A)$^+$ RNA from multiple human tissues (blot from Clontech) were probed with radioactively labelled insert of the 2.1 kb VEGF-C cDNA clone. Northern blotting and hybridization analysis showed that the 2.4 kb RNA and smaller amounts of a 2.0 kb mRNA are expressed in multiple human tissues, most prominently in the heart, placenta, muscle, ovary and small intestine (Fig. 16A). Very little VEGF-C RNA was seen in the brain, liver or thymus and

peripheral blood leukocytes (pbl) appeared negative. A similar analysis of RNA from human fetal tissues (Fig. 16B) shows that VEGF-C is highly expressed in the kidney and lung and to a lesser degree in the liver, while essentially no expression is detected in the brain. Interestingly, VEGF expression correlates with VEGF-C expression in these tissues, whereas VEGF-B is higly expressed in all tissues analysed.

EXAMPLE 17

The VEGF-C Gene Localizes To Chromosome 4q34

A DNA panel of 24 interspecies somatic cell hybrids, which had retained one or two human chromosomes, was used for the chromosomal localization of the VEGF-C gene (Bios Laboratories, Inc., New Haven, CT). Primers were designed to amplify an about 250 bp fragment of the VEGF-C gene from somatic cell hybrid DNA. The primers and conditions for polymerase chain reaction (PCR) were 5'-TGAGTGATTTGTAGCTGCTGTG-3' (forward) [SEQ ID NO:34] and 5'-TATTGCAGCAACCCCCACATCT-3' (reverse) [SEQ ID NO:35] for VEGF-C (94 C, 60s/62 C, 45s/72 C, 60s). The PCR products were evaluated by electrophoresis in 1% agarose gels and visualized by ethidium bromide staining in ultraviolet light. [α-^{32}P]-dCTP-labelled cDNA inserts of a plasmid representing the complete VEGF-C coding domain was used as a probe in Southern blotting and hybridization analysis of the somatic cell hybrid DNAs as instructed by the supplier (Bios Laboratories).

The cell lines for fluorescence *in situ* hybridization (FISH) were obtained from the American Type Culture Collection (Rockville, MD). Purified DNA from P1 clones 7660 and 7661 (VEGF-C) (Genome Systems, Inc., St. Louis, MO) were confirmed positive by Southern blotting of Eco RI-digested DNA followed by hybridization with the VEGF-C cDNA. The P1 clones were then labelled by nick translation either with biotin-11-dUTP, biotin-14-ATP (Sigma Chemical Co., St. Louis, MO) or digoxigenin 11-dUTP (Boehringer Mannheim GmbH, Mannheim, Germany) according to standard protocols. PHA-stimulated peripheral blood lymphocyte cultures were treated with 5-bromodeoxyuridine (BrdU) at an early replicating phase to induce G-banding. See Takahashi *et al.*, *Human Genet.*, *86*:14-16 (1995); Lemieux *et al.*, *Cytogenet. Cell Genet.*, *59*:311-12 (1992). The FISH procedure was carried out in 50% formamide, 10% dextran sulphate in 2x SSC using well-

known procedures. See, e.g., Rytkönnen *et al.*, *Cytogenet, Cell Genet.*, *68*:61-63 (1995); Lichter *et al.*, *Proc. Natl. Acad. Sci. USA*, *85*:9664-68 (1988). Repetitive sequences were suppressed with 50-fold excess of Cot-1 DNA (BRL, Gaithesburg, MD) compared with the labeled probe. Specific hybridization signals were detected by incubating the hybridized slides in labelled antidigoxigenin antibodies, followed by counterstaining with 0.1mmol/L 4,6-diamino-2-phenylindole. Probe detection for two-color experiments was accomplished by incubating the slides in fluorescein isothiocyanate (FITC)-anti-digoxigenin antibodies (Sigma Chemical Co.) and Texas red-avidin (Vector Laboratories, Burlingame, CA) or rhodamine-anti-digoxigenin and FITC-avidin.

Multi-color digital image analysis was used for acquisition, display and quantification of hybridization signals of metaphase chromosomes. The system contains a PXL camera (Photometrics Inc., Tucson, AZ) attached to a PowerMac 7100/Av workstation. IPLab software controls the camera operation, image acquisition and Ludl Filter wheel. At least 50 nuclei were scored. Overlapping nuclei and clusters of cells were ignored. A slide containing normal lymphocyte metaphase spreads and interphase nuclei was included in each experiment to control for the efficiency and specificity of the hybridization.

In order to determine the chromosomal localization of the human VEGF-C gene, DNAs from human rodent somatic cell hybrids containing defined sets of human chromosomes were analysed by Southern blotting and hybridization with the VEGF-C cDNA probe. Among 24 DNA samples on the hybrid panel, representing different human chromosomes, human-specific signals were observed only in hybrids which contained human chromosome 4. The results were confirmed by PCR of somatic cell hybrid DNA using VEGF-C specific primers, where amplified bands were obtained only from DNAs containing human chromosome 4.

A genomic P1 plasmid for VEGF-C was isolated using specific primers and PCR and verified by Southern blotting and hybridization using a VEGF-C specific cDNA probe. The chromosomal localization of VEGF-C was further studied using metaphase FISH. Using the P1 probe for VEGF-C in FISH a specific hybridization to the 4q34 chromosomal band was detected in 40 out of 44 metaphases (Fig. 17). Double-fluorochrome hybridization using

a cosmid probe specific for the aspartylglucosaminidase (AGA) gene showed that VEGF-C is located just proximal to the AGA gene previously mapped to the 4q34-35 chromosomal band.

Biotin labelled VEGF-C P1 and digoxigenin labeled AGA cosmid probes were hybridized simultaneously to metaphase chromosomes. This experiment demonstrated that the AGA gene is more telomerically located than the VEGF-C gene. The foregoing example demonstrates the utility of polynucleotides of the invention as chromosomal markers.

EXAMPLE 18

Effect of glucose concentration and hypoxia on VEGF, VEGF-B and VEGF-C mRNA levels in c6 glioblastoma cells

Confluent cultures of C6 cells (ATCC CCL 107) were grown on 10 cm diameter tissue culture plates containing 2.5 ml of DMEM and 5% fetal calf serum plus antibiotics. The cultures were exposed for 16 hours to normoxia in a normal cell culture incubator containing 5% CO_2 (Fig. 18: lanes marked -) or hypoxia (Fig. 18: lanes marked +) by closing the culture plates in an airtight glass chamber and burning a piece of wood inside until the flame was extinguished due to lack of oxygen. Polyadenylated RNA was isolated (as in the other examples), and 8 micrograms of the RNA was electrophoresed and blot-hybridized with a mixture of the VEGF, VEGF-B and VEGF-C probes (see Fig. 12). The results show that hypoxia strongly induces VEGF (VEGF-A) mRNA expression (compare lanes - and +), both in low and high glucose, but has no significant effect on the VEGF-B mRNA levels. The VEGF-C mRNA isolated from hypoxic cells runs slightly faster in gel electrophoresis and an extra band of faster mobility can be seen below the upper mRNA band. This observation suggests that hypoxia affects VEGF-C RNA processing. One explanation for this observation is that VEGF-C mRNA splicing is altered, affecting the VEGF-C open reading frame and resulting in an alternative VEGF-C protein being produced by hypoxic cells. Such alternative forms of VEGF-C and VEGF-C-encoding polynucleotides are contemplated as an aspect of the invention. This data indicates screening and diagnostic utilities for polynucleotides and polypeptides of the invention, such as methods whereby a biological sample is screened for the hypoxia-induced

form of VEGF-C and/or VEGF-C mRNA. The data further suggests a therapeutic indication for antibodies and/or other inhibitors of the hypoxia-induced form of VEGF-C or the normal form of VEGF-C.

EXAMPLE 19

Pulse-chase labeling and immunoprecipitation of VEGF-C polypeptides from 293 cells transfected with VEGF-C expression vector.

The following VEGF-C branched amino-terminal peptide, designated PAM126, was synthesized for production of anti-VEGF-C antiserum:

NH_2-E-E-T-I-K-F-A-A-A-H-Y-N-T-E-I-L-K-COOH (SEQ ID NO: 39).

In particular, PAM126 was synthesized as a branched polylysine structure K3PA4 having four peptide acid (PA) chains attached to two available lysine (K) residues. The synthesis was performed on a 433A Peptide Synthesizer (Applied Biosystems) using Fmoc-chemistry and TentaGel S MAP RAM10 resin mix (RAPP Polymere GmbH, Tubingen, Germany), yielding both cleavable and resin-bound peptides. The cleavable peptide was purified via reverse phase HPLC and was used together with the resin-bound peptide in immunizations. The correctness of the synthesis products were confirmed using mass-spectroscopy (Lasermatt).

The peptide was dissolved in phosphate buffered saline (PBS), mixed with Freund's adjuvant, and used for immunization of rabbits at bi-weekly intervals using methods standard in the art (Harlow and Lane, *Antibodies, a laboratory manual*, Cold Spring Harbor Laboratory Press (1988)). Antisera obtained after the fourth booster immunization was used for immunoprecipitation of VEGF-C in pulse-chase experiments, as described below.

For pulse-chase analysis, 293 cells transfected with a VEGF-C expression vector (i.e., the FLT4-L cDNA inserted into the pREP7 expression vector as described above) were incubated for 30 minutes in methionine-free, cysteine-free, serum-free DMEM culture medium at 37 C. The medium was then changed, and 200 microCuries of ^{35}S-methionine and cysteine (Promix, Amersham, Buckinghamshire, England) was added. The cell layers were incubated in this labeling medium for two hours, washed with PBS, and

incubated for 0, 15, 30, 60, 90, 120, or 180 minutes in serum-free DMEM (chase). After the various chase periods, the medium was collected, the cells were again washed two times in PBS, and lysed in immunoprecipitation buffer. The VEGF-C polypeptides were analyzed from both the culture medium and from the cell lysates by immunoprecipitation, using the VEGF-C-specific antiserum raised against the NH_2-terminal peptide (PAM126) of the 23 kD VEGF-C form. Immunoprecipitated polypeptides were analyzed via SDS-PAGE followed by autoradiography.

Referring to Fig. 19, the resultant autoradiograms demonstrate that immediately after a 2 hour labeling (chase time 0), the VEGF-C vector-transfected cells contained a radioactive 55 kD polypeptide band, which is not seen in mock-transfected cells (M). This 55 kD polypeptide band gradually diminishes in intensity with increasing chase periods, and is no longer detected in the cells by 180 minutes of chase. A 32 kD polypeptide band also is observed in VEGF-C transfected cells (and not mock-transfected cells). This 32 kD band disappears with similar kinetics to that of the 55 kD band. Simultaneously, increasing amounts of 32 kD (arrow) and subsequently 23 kD (arrow) and 14 kD polypeptides appear in the medium.

Collectively, the data from the pulse-chase experiments indicate that the 55 kD intracellular polypeptide represents a pro-VEGF-C polypeptide, which is not secreted from cells, but rather is first proteolytically cleaved into the 32 kD form. The 32 kD form is secreted and simultaneously further processed by proteolysis into the 23 kD and 14 kD forms. Without intending to be limited to a particular theory, it is believed that processing of the VEGF-C precursor occurs as removal of a signal sequence, removal of the COOH-terminal domain (BRP3), and removal of an amino terminal peptide, resulting in a VEGF-C polypeptide having the TEE... amino terminus.

At high resolution, the 23 kD polypeptide band appears as a closely spaced polypeptide doublet, suggesting heterogeneity in cleavage or glycosylation.

EXAMPLE 20

Isolation of Mouse cDNA Clones Encoding VEGF-C

To clone a mouse variant of VEGF-C, approximately 1 x 10^6 bacteriophage lambda clones of a commercially-available 12 day mouse

embryonal cDNA library (lambda EXlox library, Novagen # 69632-1) were screened with a radiolabeled fragment of human VEGF-C cDNA containing nucleotides 495 to 1661 of SEQ ID NO:32. One positive clone was isolated.

A 1323 bp EcoRI/HindIII fragment of the insert of the isolated mouse cDNA clone was subcloned into the corresponding sites of the pBluescript SK+ vector (Stratagene) and sequenced. The cDNA sequence of this clone was homologous to the human VEGF-C sequence reported herein, except that about 710 bp of 5'-end sequence present in the human clone was not present in the mouse clone.

For further screening of mouse cDNA libraries, a HindIII-BstXI (Hind III site is from the pBluescript SK+ polylinker) fragment of 881 bp from the coding region of the mouse cDNA clone was radiolabeled and used as a probe to screen two additional mouse cDNA libraries. Two additional cDNA clones from an adult mouse heart ZAP II cDNA library (Stratagene # 936306) were identified. Three additional clones also were isolated from a mouse heart 5'-stretch-plus cDNA library in λgt11 (Clontech #ML5002b). Of the latter three clones, one was found to contain an insert of about 1.9 kb. The insert of this cDNA clone was subcloned into EcoRI sites of pBluescript SK+ vector and both strands of this clone were completely sequenced, resulting in the nucleotide and deduced amino acid sequences shown in SEQ ID NOs: 40 and 41.

It is contemplated that the polypeptide corresponding to SEQ ID NO: 41 is processed into a mature mouse VEGF-C protein, in a manner analogous to the processing of the human VEGF-C prepropeptide. Putative cleavage sites for the mouse protein are identified using procedures outlined above for identification of cleavage sites for the human VEGF-C polypeptide.

The foregoing example demonstrates the utility of polynucleotides of the invention for identifying and isolating polynucleotides encoding other non-human mammalian variants of VEGF-C. Such identified and isolated polynucleotides, in turn, can be expressed (using procedures similar to those described in preceding examples) to produce recombinant polypeptides corresponding to non-human mammalian variants of VEGF-C.

Deposit of Biological Materials: Plasmid FLT4-L has been deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Dr., Rockville MD 20952 (USA), pursuant to the provisions of the Budapest Treaty, and has been assigned a deposit date of 24 July 1995 and ATCC accession number 97231.

While the present invention has been described in terms of specific embodiments, it is understood that variations and modifications will occur to those in the art. Accordingly, only such limitations as appear in the appended claims should be placed on the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Alitalo, Kari
Joukov, Vladimir

(ii) TITLE OF INVENTION: Receptor Ligand

(iii) NUMBER OF SEQUENCES: 41

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
(B) STREET: 6300 Sears Tower, 233 South Wacker Drive
(C) CITY: Chicago
(D) STATE: Illinois
(E) COUNTRY: United States of America
(F) ZIP: 60606-6402

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Gass, David A.
(B) REGISTRATION NUMBER: 38,153
(C) REFERENCE/DOCKET NUMBER: 28113/33118

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 312/474-6300
(B) TELEFAX: 312/474-0448
(C) TELEX: 25-3856

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGTCCTCGCT GTCCTTGTCT 20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACATGCATGC CACCATGCAG CGGGGCGCCG CGCTGTCCCT GCGACTGTGG CTCTGCCTGG 60

GACTCCTGGA 70

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACATGCATGC CCCGCCGGTC ATCC 24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGAATTCCC CATGACCCCA AC 22

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCATCGATGG ATCCTACCTG AAGCCGCTTT CTT 33

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATTTAGGTGA CACTATA 17

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCATCGATGG ATCCCGATGC TGCTTAGTAG CTGT 34

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Pro Met Thr Pro Thr Thr Tyr Lys Gly Ser Val Asp Asn Gln Thr Asp
1 5 10 15

Ser Gly Met Val Leu Ala Ser Glu Glu Phe Glu Gln Ile Glu Ser Arg
20 25 30

His Arg Gln Glu Ser Gly Phe Arg
35 40

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGGAGTCGA CTTGGCGGAC T 21

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCGGATCCC TAGTGATGGT GATGGTGATG TCTACCTTCG ATCATGCTGC CCTTATCCTC 60

) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linea-

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCCAAGCTTG GATCCAAGTG GCTACTCCAT GACC 34

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTTGCCTGTG ATGTGCACCA 20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Xaa Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr Asn Thr Glu Ile
1 5 10 15

Leu Lys

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCAGARGARA CNATHAA 17

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Glu Glu Thr Ile Lys
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAYTTNARD ATYTCNGT 18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Thr Glu Ile Leu Lys
1 5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCTGCA GCACACTACA AC 22

(2) INFORMATION FOR SEQ ID NO:19:

(i) SSQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCNGTGTTGT AGTGTGCTG 19

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STPANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Ala Ala His Tyr Asn Thr Glu
1 5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TAATACGACT CACTATAGGG 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTTGTAGTGT GCTGCAGCGA ATTT 24

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Lys Phe Ala Ala Ala His Tyr Asn
1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TCACTATAGG GAGACCCAAG C 21

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 219 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TCACTATAGG GAGACCCAAG CTTGGTACCG AGCTCGGATC CACTAGTAAC GGCCGCCAGT 60

GTGGTGGAAT TCGACGAACT CATGACTGTA CTCTACCCAG AATATTGGAA AATGTACAAG 120

TGTCAGCTAA GGCAAGGAGG CTGGCAACAT AACAGAGAAC AGGCCAACCT CAACTCAAGG 180

ACAGAAGAGA CTATAAAATT CGCTGCAGCA CACTACAAC 219

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ACAGAGAACA GGCCAACC 18

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCTAGCATTT AGGTGACAC 19

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAGAGACTAT AAAATTCGCT GCAGC 25

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCCTCTAGAT GCATGCTCGA 20

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTTGTAGTGT GCTGCAGCGA ATTT 24

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCACTATAGG GAGACCCAAG C 21

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1997 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 352..1608

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 658..1608

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
CCCGCCCCGC CTCTCCAAAA AGCTACACCG ACGCGGACCG CGGCGGCGTC CTCCCTCGCC      60

CTCGCTTCAC CTCGCGGGCT CCGAATGCGG GGAGCTCGGA TGTCCGGTTT CCTGTGAGGC     120

TTTTACCTGA CACCCGCCGC CTTTCCCCGG CACTGGCTGG GAGGGCCCCC TGCAAAGTTG     180

GGAACGCGGA GCCCCGGACC CGCTCCCGCC GCCTCCGGCT CGCCCAGGGG GGGTCGCCGG     240

GAGGAGCCCG GGGGAGAGGG ACCAGGAGGG GCCCGCGGCC TCGCAGGGGC GCCCGCGCCC     300

CCACCCCTGC CCCCGCCAGC GGACCGGTCC CCCACCCCCG GTCCTTCCAC C ATG CAC      357
                                                          Met His
                                                          -102

TTG CTG GGC TTC TTC TCT GTG GCG TGT TCT CTG CTC GCC GCT GCG CTG      405
Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala Ala Leu
-100                 -95                 -90                 -85

CTC CCG GGT CCT CGC GAG GCG CCC GCC GCC GCC GCC GCC TTC GAG TCC      453
Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Ala Phe Glu Ser
                -80                 -75                 -70

GGA CTC GAC CTC TCG GAC GCG GAG CCC GAC GCG GGC GAG GCC ACG GCT      501
Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala Thr Ala
            -65                 -60                 -55

TAT GCA AGC AAA GAT CTG GAG GAG CAG TTA CGG TCT GTG TCC AGT GTA      549
Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser Ser Val
        -50                 -45                 -40

GAT GAA CTC ATG ACT GTA CTC TAC CCA GAA TAT TGG AAA ATG TAC AAG      597
Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met Tyr Lys
    -35                 -30                 -25

TGT CAG CTA AGG AAA GGA GGC TGG CAA CAT AAC AGA GAA CAG GCC AAC      645
Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln Ala Asn
-20                 -15                 -10                  -5

CTC AAC TCA AGG ACA GAA GAG ACT ATA AAA TTT GCT GCA GCA CAT TAT      693
Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr
                1                   5                  10

AAT ACA GAG ATC TTG AAA AGT ATT GAT AAT GAG TGG AGA AAG ACT CAA      741
Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln
            15                  20                  25

TTC ATG CCA CGG GAG GTG TGT ATA GAT GTG GGG AAG GAG TTT GGA GTC      789
Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Val
        30                  35                  40

GCG ACA AAC ACC TTC TTT AAA CCT CCA TGT GTG TCC GTC TAC AGA TGT      837
Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys
    45                  50                  55                  60

GGG GGT TGC TGC AAT AGT GAG GGG CTG CAG TGC ATG AAC ACC AGC ACG      885
Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr Ser Thr
                65                  70                  75
```

```
AGC TAC CTC AGC AAG ACG TTA TTT GAA ATT ACA GTG CCT CTC TCT CAA          933
Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln
                80                  85                  90

GGC CCC AAA CCA GTA ACA ATC AGT TTT GCC AAT CAC ACT TCC TGC CGA          981
Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg
            95                  100                 105

TGC ATG TCT AAA CTG GAT GTT TAC AGA CAA GTT CAT TCC ATT ATT AGA         1029
Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg
        110                 115                 120

CGT TCC CTG CCA GCA ACA CTA CCA CAG TGT CAG GCA GCG AAC AAG ACC         1077
Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr
    125                 130                 135                 140

TGC CCC ACC AAT TAC ATG TGG AAT AAT CAC ATC TGC AGA TGC CTG GCT         1125
Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys Leu Ala
                145                 150                 155

CAG GAA GAT TTT ATG TTT TCC TCG GAT GCT GCA GAT GAC TCA ACA GAT         1173
Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser Thr Asp
            160                 165                 170

GGA TTC CAT GAC ATC TGT GGA CCA AAC AAG GAG CTG GAT GAA GAG ACC         1221
Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu Glu Thr
        175                 180                 185

TGT CAG TGT GTC TGC AGA GCG GGG CTT CGG CCT GCC AGC TGT GGA CCC         1269
Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys Gly Pro
    190                 195                 200

CAC AAA GAA CTA GAC AGA AAC TCA TGC CAG TGT GTC TGT AAA AAC AAA         1317
His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys Asn Lys
205                 210                 215                 220

CTC TTC CCC AGC CAA TGT GGG GCC AAC CGA GAA TTT GAT GAA AAC ACA         1365
Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr
                225                 230                 235

TGC CAG TGT GTA TGT AAA AGA ACC TGC CCC AGA AAT CAA CCC CTA AAT         1413
Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn
            240                 245                 250

CCT GGA AAA TGT GCC TGT GAA TGT ACA GAA AGT CCA CAG AAA TGC TTG         1461
Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys Cys Leu
        255                 260                 265

TTA AAA GGA AAG AAG TTC CAC CAC CAA ACA TGC AGC TGT TAC AGA CGG         1509
Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg
    270                 275                 280

CCA TGT ATG AAC CGC CAG AAG GCT TGT GAG CCA GGA TTT TCA TAT AGT         1557
Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser Tyr Ser
285                 290                 295                 300

GAA CAA GTG TGT CGT TGT GTC CCT TCA TAT TGG AAA AGA CCA CAA ATG         1605
Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro Gln Met
                305                 310                 315

AGC TAAGATTGTA CTGTTTTCCA GTTCATCGAT TTTCTATTAT GGAAAACTGT               1658
Ser

GTTGCCACAG TAGAACTGTC TGTGAACAGA GAGACCCTTG TCGGTCCATG CTAACAAAGA       1718

CAAAAGTCTG TCTTTCCTGA ACCATGTGGA TAACTTTACA GAAATGGACT GGAGCTCATC       1778

TGCAAAAGGC CTCTTGTAAA GACTGGTTTT CTGCCAATGA CCAAACAGCC AAGATTTTCC       1838

TCTTGTGATT TCTTTAAAAG AATGACTATA TAATTTATTT CCACTAAAAA TATTGTTTCT       1898

GCATTCATTT TTATAGCAAC AACAATTGGT AAAACTCACT GTGATCAATA TTTTTATATC       1958
```

ATGCAAAATA TGTTTAAAAT AAAATGAAAA TTGTATTAT 1997

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 419 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Met His Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala
-102      -100                 -95                 -90

Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Ala Phe
     -85                 -80                 -75

Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala
-70                 -65                 -60                 -55

Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser
                -50                 -45                 -40

Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met
            -35                 -30                 -25

Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln
        -20                 -15                 -10

Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala
    -5                  1                   5                  10

His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys
                15                  20                  25

Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe
            30                  35                  40

Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr
        45                  50                  55

Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr
    60                  65                  70

Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu
75                  80                  85                  90

Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser
                95                  100                 105

Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile
            110                 115                 120

Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn
        125                 130                 135

Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys
    140                 145                 150

Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser
155                 160                 165                 170

Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu
                175                 180                 185

Glu Thr Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys
            190                 195                 200

Gly Pro His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys
        205                 210                 215
```

Asn Lys Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu
220 225 230

Asn Thr Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro
235 240 245 250

Leu Asn Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys
255 260 265

Cys Leu Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr
270 275 280

Arg Arg Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser
285 290 295

Tyr Ser Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro
300 305 310

Gln Met Ser
315

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TGAGTGATTTGTAGCTGCTGTG 22

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TATTGCAGCAACCCCCACATCT 22

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4416 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CCACGCGCAG CGGCCGGAGA TGCAGCGGGG CGCCGCGCTG TGCCTGCGAC TGTGGCTCTG 60

CCTGGGACTC CTGGACGGCC TGGTGAGTGG CTACTCCATG ACCCCCCCGA CCTTGAACAT 120

CACGGAGGAG TCACACGTCA TCGACACCGG TGACAGCCTG TCCATCTCCT GCAGGGGACA 180

GCACCCCCTC GAGTGGGCTT GGCCAGGACC TCAGGAGGCG CCAGCCACCG GAGACAAGGA 240

CAGCGAGGAC ACGGGGGTGG TGCGAGACTG CGAGGGCACA GACGCCAGGC CCTACTGCAA 300

GGTGTTGCTG CTGCACGAGG TACATGCCAA CGACACAGGC AGCTACGTCT GCTACTACAA 360

GTACATCAAG GCACGCATCG AGGGCACCAC GGCCGCCAGC TCCTACGTGT TCGTGAGAGA 420

CTTTGAGCAG CCATTCATCA ACAAGCCTGA CACGCTCTTG GTCAACAGGA AGGACGCCAT 480

GTGGGTGCCC TGTCTGGTGT CCATCCCCGG CCTCAATGTC ACGCTGCGCT CGCAAAGCTC 540

GGTGCTGTGG CCAGACGGGC AGGAGGTGGT GTGGGATGAC CGGCGGGGCA TGCTCGTGTC 600

CACGCCACTG CTGCACGATG CCCTGTACCT GCAGTGCGAG ACCACCTGGG GAGACCAGGA 660

CTTCCTTTCC AACCCCTTCC TGGTGCACAT CACAGGCAAC GAGCTCTATG ACATCCAGCT 720

GTTGCCCAGG AAGTCGCTGG AGCTGCTGGT AGGGGAGAAG CTGGTCCTGA ACTGCACCGT 780

GTGGGCTGAG TTTAACTCAG GTGTCACCTT TGACTGGGAC TACCCAGGGA AGCAGGCAGA 840

GCGGGGTAAG TGGGTGCCCG AGCGACGCTC CCAGCAGACC CACACAGAAC TCTCCAGCAT 900

CCTGACCATC CACAACGTCA GCCAGCACGA CCTGGGCTCG TATGTGTGCA AGGCCAACAA 960

CGGCATCCAG CGATTTCGGG AGAGCACCGA GGTCATTGTG CATGAAAATC CCTTCATCAG 1020

CGTCGAGTGG CTCAAAGGAC CCATCCTGGA GGCCACGGCA GGAGACGAGC TGGTGAAGCT 1080

GCCCGTGAAG CTGGCAGCGT ACCCCCCGCC CGAGTTCCAG TGGTACAAGG ATGGAAAGGC 1140

ACTGTCCGGG CGCCACAGTC CACATGCCCT GGTGCTCAAG GAGGTGACAG AGGCCAGCAC 1200

AGGCACCTAC ACCCTCGCCC TGTGGAACTC CGCTGCTGGC CTGAGGCGCA ACATCAGCGT 1260

GGAGCTGGTG GTGAATGTGC CCCCCCAGAT ACATGAGAAG GAGGCCTCCT CCCCCAGCAT 1320

CTACTCGCGT CACAGCCGCC AGGCCCTCAC CTGCACGGCC TACGGGGTGC CCCTGCCTCT 1380

CAGCATCCAG TGGCACTGGC GGCCCTGGAC ACCCTGCAAG ATGTTTGCCC AGCGTAGTCT 1440

CCGGCGGCGG CAGCAGCAAG ACCTCATGCC ACAGTGCCGT GACTGGAGGG CGGTGACCAC 1500

GCAGGATGCC GTGAACCCCA TCGAGAGCCT GGACACCTGG ACCGAGTTTG TGGAGGGAAA 1560

GAATAAGACT GTGAGCAAGC TGGTGATCCA GAATGCCAAC GTGTCTGCCA TGTACAAGTG 1620

TGTGGTCTCC AACAAGGTGG GCCAGGATGA GCGGCTCATC TACTTCTATG TGACCACCAT 1680

CCCCGACGGC TTCACCATCG AATCCAAGCC ATCCGAGGAG CTACTAGAGG GCCAGCCGGT 1740

GCTCCTGAGC TGCCAAGCCG ACAGCTACAA GTACGAGCAT CTGCGCTGGT ACCGCCTCAA 1800

CCTGTCCACG CTGCACGATG CGCAGGGGAA CCCGCTTCTG CTCGACTGCA AGAACGTGCA 1860

TCTGTTCGCC ACCCCTCTGG CCGCCAGCCT GGAGGAGGTG GCACCTGGGG CGCGCCACGC 1920

CACGCTCAGC CTGAGTATCC CCCGCGTCGC GCCCGAGCAC GAGGGCCACT ATGTGTGCGA 1980

AGTGCAAGAC CGGCGCAGCC ATGACAAGCA CTGCCACAAG AAGTACCTGT CGGTGCAGGC 2040

CCTGGAAGCC CCTCGGCTCA CGCAGAACTT GACCGACCTC CTGGTGAACG TGAGCGACTC 2100

GCTGGAGATG CAGTGCTTGG TGGCCGGAGC GCACGCGCCC AGCATCGTGT GGTACAAAGA 2160

CGAGAGGCTG CTGGAGGAAA AGTCTGGAGT CGACTTGGCG GACTCCAACC AGAAGCTGAG 2220

CATCCAGCGC GTGCGCGAGG AGGATGCGGG ACGCTATCTG TGCAGCGTGT GCAACGCCAA 2280

GGGCTGCGTC AACTCCTCCG CCAGCGTGGC CGTGGAAGGC TCCGAGGATA AGGGCAGCAT 2340

GGAGATCGTG ATCCTTGTCG GTACCGGCGT CATCGCTGTC TTCTTCTGGG TCCTCCTCCT 2400

CCTCATCTTC TGTAACATGA GGAGGCCGGC CCACGCAGAC ATCAAGACGG GCTACCTGTC 2460

CATCATCATG GACCCCGGGG AGGTGCCTCT GGAGGAGCAA TGCGAATACC TGTCCTACGA 2520

TGCCAGCCAG TGGGAATTCC CCCGAGAGCG GCTGCACCTG GGGAGAGTGC TCGGCTACGG 2580

CGCCTTCGGG AAGGTGGTGC AGCCTCCGC TTTCGGCATC CACAAGGGCA GCAGCTGTGA 2640

CACCGTGGCC GTGAAAATGC TGAAAGAGGG CGCCACGGCC AGCGAGCACC GCGCGCTGAT 2700

GTCGGAGCTC AAGATCCTCA TTCATCGG CAACCACCTC AACGTGGTCA ACCTCCTCGG 2760

GGCGTGCACC AAGCCGCAGG GCCCCCTCAT GGTGATCGTG GAGTTCTGCA AGTACGGCAA 2820

CCTCTCCAAC TTCCTGCGCG CCAAGCGGG. CGCCTTCAGC CCCTGCGCGG AGAAGTCTCC 2880

CGAGCAGCGC GGACGCTTCC GCGCCATGGT GGAGCTCGCC AGGCTGGATC GGAGGCGGCC 2940

GGGGAGCAGC GACAGGGTCC TCTTCGCGCG GTTCTCGAAG ACCGAGGGCG GAGCGAGGCG 3000

GGCTTCTCCA GACCAAGAAG CTGAGGACCT GTGGCTGAGC CCGCTGACCA TGGAAGATCT 3060

TGTCTGCTAC AGCTTCCAGG TGGCCAGAGG GATGGAGTTC CTGGCTTCCC GAAAGTGCAT 3120

CCACAGAGAC CTGGCTGCTC GGAACATTCT GCTGTCGGAA AGCGACGTGG TGAAGATCTG 3180

TGACTTTGGC CTTGCCCGGG ACATCTACAA AGACCCTGAC TACGTCCGCA AGGGCAGTGC 3240

CCGGCTGCCC CTGAAGTGGA TGGCCCCTGA AAGCATCTTC GACAAGGTGT ACACCACGCA 3300

GAGTGACGTG TGGTCCTTTG GGGTGCTTCT CTGGGAGATC TTCTCTCTGG GGGCCTCCCC 3360

GTACCCTGGC GTGCAGATCA ATGAGGAGTT CTGCCAGCGG CTGAGAGACG GCACAAGGAT 3420

GAGGGCCCCG GAGCTGGCCA CTCCCGCCAT ACGCCGCATC ATGCTGAACT GCTGGTCCGG 3480

AGACCCCAAG GCGAGACCTG CATTCTCGGA GCTGGTGGAG ATCCTGGGGG ACCTGCTCCA 3540

GGGCAGGGGC CTGCAAGAGG AAGAGGAGGT CTGCATGGCC CCGCGCAGCT CTCAGAGCTC 3600

AGAAGAGGGC AGCTTCTCGC AGGTGTCCAC CATGGCCCTA CACATCGCCC AGGCTGACGC 3660

TGAGGACAGC CCGCCAAGCC TGCAGCGCCA CAGCCTGGCC GCCAGGTATT ACAACTGGGT 3720

GTCCTTTCCC GGGTGCCTGG CCAGAGGGGC TGAGACCCGT GGTTCCTCCA GGATGAAGAC 3780

ATTTGAGGAA TTCCCCATGA CCCCAACGAC CTACAAAGGC TCTGTGGACA ACCAGACAGA 3840

CAGTGGGATG GTGCTGGCCT CGGAGGAGTT TGAGCAGATA GAGAGCAGGC ATAGACAAGA 3900

AAGCGGCTTC AGCTAGCTGA AGCAGAGAGA GAGAAGGCAG CATACGTCAG CATTTTCTTC 3960

TCTGCACTTA TATGAAAGAT CAAAGACTTT AAGACTTTCG CTATTTCTTC TACTGCTATC 4020

TACTACAAAC TTCAAAGAGG AACCAGGAGG ACAAGAGGAG CATGAAAGTG GACAAGGAGT 4080

GTGACCACTG AAGCACCACA GGGAAGGGGT TAGGCCTCCG GATGACTGCG GGCAGGCCTG 4140

GATAATATCC AGCCTCCCAC AAGAAGCTGG TGGAGCAGAG TGTTCCCTGA CTCCTCCAAG 4200

GAAAGGGAGA CGCCCTTTCA TGGTCTGCTG AGTAACAGGT GCNTTCCCAG ACACTGGCGT 4260

TACTGCTTGA CCAAAGAGCC CTCAAGCGGC CCTTATGCCA GCGTGACAGA GGGCTCACCT 4320

CTTGCCTTCT AGGTCACTTC TCACACAATG TCCCTTCAGC ACCTGACCCT GTGCCCGCCA 4380

GTTATTCCTT GGTAATATGA GTAATACATC AAAGAG 4416

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4273 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AAGCTTATCG ATTTCGAACC CGGGGGTACC GAATTCCTCG AGTCTAGAGG AGCATGCCTG 60

CAGGTCGACC GGGCTCGATC CCCTCGCGAG TTGGTTCAGC TGCTGCCTGA GGCTGGACGA 120

CCTCGCGGAG TTCTACCGGC AGTGCAAATC CGTCGGCATC CAGGAAACCA GCAGCGGCTA 180

TCCGCGCATC CATGCCCCCG AACTGCAGGA GTGGGGAGGC ACGATGGCCG CTTTGGTCCC 240

GGATCTTTGT GAAGGAACCT TACTTCTGTG GTGTGACATA ATTGGACAAA CTACCTACAG 300

AGATTTAAAG CTCTAAGGTA AATATAAAAT TTTTAAGTGT ATAATGTGTT AAACTACTGA 360

TTCTAATTGT TTGTGTATTT TAGATTCCAA CCTATGGAAC TGATGAATGG GAGCAGTGGT 420

GGAATGCCTT TAATGAGGAA AACCTGTTTT GCTCAGAAGA AATGCCATCT AGTGATGATG 480

AGGCTACTGC TGACTCTCAA CATTCTACTC CTCCAAAAAA GAAGAGAAAG GTAGAAGACC 540

CCAAGGACTT TCCTTCAGAA TTGCTAAGTT TTTTGAGTCA TGCTGTGTTT AGTAATAGAA 600

CTCTTGCTTG CTTTGCTATT TACACCACAA AGGAAAAAGC TGCACTGCTA TACAAGAAAA 660

TTATGGAAAA ATATTCTGTA ACCTTTATAA GTAGGCATAA CAGTTATAAT CATAACATAC 720

TGTTTTTTCT TACTCCACAC AGGCATAGAG TGTCTGCTAT TAATAACTAT GCTCAAAAAT 780

TGTGTACCTT TAGCTTTTTA ATTTGTAAAG GGGTTAATAA GGAATATTTG ATGTATAGTG 840

CCTTGACTAG AGATCATAAT CAGCCATACC ACATTTGTAG AGGTTTTACT TGCTTTAAAA 900

AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT TGTTGTTAAC 960

TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT 1020

AAAGCATTTT TTTCACTGCA TTCTAGTTGT GGTTTGTCCA AACTCATCAA TGTATCTTAT 1080

CATGTCTGGA TCTGCCGGTC TCCCTATAGT GAGTCGTATT AATTTCGATA AGCCAGGTTA 1140

ACCTGCATTA ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT GGGCGCTCTT 1200

CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA GCGGTATCAG 1260

CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA 1320

TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGACGCGTTG CTGGCGTTTT 1380

TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC 1440

GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT 1500

CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG 1560

TGGCGCTTTC TCAATGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA 1620

AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT 1680

ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA 1740

ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA 1800

ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG CCAGTTACCT 1860

TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT 1920

TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA GATCCTTTGA 1980

TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC ACGTTAAGGG ATTTTGGTCA 2040

TGAGATTATC AAAAAGGATC TTCACCTAGA TCCTTTTAAA TTAAAAATGA AGT .TTAAAT 2100
CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA CCAATGCTTA ATCAGTGAGG 2160
CACCTATCTC AGCGATCTGT CTATTTCGTT CATCCATAGT TGCCTGACTC CCCGTCGTGT 2220
AGATAACTAC GATACGGGAG GGCTTACCAT CTGGCCCCAG TGCTGCAATG ATACCGCGAG 2280
ACCCACGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA GCCAGCCGGA AGGGCCGAGC 2340
GCAGAAGTGG TCCTGCAACT TTATCCGCCT CCATCCAGTC TATTAATTGT TGCCGGGAAG 2400
CTAGAGTAAG TAGTTCGCCA GTTAATAGTT TGCGCAACGT TGTTGCCATT GCTACAGGCA 2460
TCGTGGTGTC ACGCTCGTCG TTTGGTATGG CTTCATTCAG CTCCGGTTCC CAACGATCAA 2520
GGCGAGTTAC ATGATCCCCC ATGTTGTGCA A AAAAGCGGT TAGCTCCTTC GGTCCTCCGA 2580
TCGTTGTCAG AAGTAAGTTG GCCGCAGTGT TATCACTCAT GGTTATGGCA GCACTGCATA 2640
ATTCTCTTAC TGTCATGCCA TCCGTAAGAT GCTTTTCTGT GACTGGTGAG TACTCAACCA 2700
AGT ATTCTG AGAATAGTGT ATGCGGCGAC CGAGTTGCTC TTGCCCGGCG TCAATACGGG 2760
ATAATACCGC GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA CGTTCTTCGG 2820
GGCGAAAACT CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA CCCACTCGTG 2880
CACCCAACTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT TTCTGGGTGA GCAAAAACAG 2940
GAAGGCAAAA TGCCGCAAAA AAGGGAATAA GGGCGACACG GAAATGTTGA ATACTCATAC 3000
TCTTCCTTTT TCAATATTAT TGAAGCATTT ATCAGGGTTA TTGTCTCATG AGCGGATACA 3060
TATTTGAATG TATTTAGAAA AATAAACAAA TAGGGGTTCC GCGCACATTT CCCCGAAAAG 3120
TGCCACCTGA CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA AATAGGCGTA 3180
TCACGAGGCC CTTTCGTCTC GCGCGTTTCG GTGATGACGG TGAAAACCTC TGACACATGC 3240
AGCTCCCGGA GACGGTCACA GCTTGTCTGT AAGCGGATGC CGGGAGCAGA CAAGCCCGTC 3300
AGGGCGCGTC AGCGGGTGTT GGCGGGTGTC GGGGCTGGCT TAACTATGCG GCATCAGAGC 3360
AGATTGTACT GAGAGTGCAC CATATGGACA TATTGTCGTT AGAACGCGGC TACAATTAAT 3420
ACATAACCTT ATGTATCATA CACATACGAT TTAGGTGACA CTATAGAACT CGAGCAGAGC 3480
TTCCAAATTG AGAGAGAGGC TTAATCAGAG ACAGAAACTG TTTGAGTCAA CTCAAGGATG 3540
GTTTGAGGGA CTGTTTAACA GATCCCCTTG GTTTACCACC TTGATATCTA CCATTATGGG 3600
ACCCCTCATT GTACTCCTAA TGATTTTGCT CTTCGGACCC TGCATTCTTA ATCGATTAGT 3660
CCAATTTGTT AAAGACAGGA TATCAGTGGT CCAGGCTCTA GTTTTGACTC AACA.TATCA 3720
CCAGCTGAAG CCTATAGAGT ACGAGCCATA GATAAAATAA AAGATTTTAT TTAGTCTCCA 3780
GAAAAAGGGG GGAATGAAAG ACCCCACCTG TAGGTTTGGC AAGCTAGCTT AAGTAACGCC 3840
ATTTTGCAAG GCATGGAAAA ATACATAACT GAGAATAGAG AAGTTCAGAT CAAGGTCAGG 3900
AACAGATGGA ACAGCTGAAT ATGGGCCAAA CAGGATATCT GTGGTAAGCA GTTCCTGCCC 3960
CGGCTCAGGG CCAAGAACAG ATGGAACAGC TGAATATGGG CCAAACAGGA TATCTGTGGT 4020
AAGCAGTTCC TGCCCCGGCT CAGGGCCAAG AACAGATGGT CCCCAGATGC GGTCCAGCCC 4080
TCAGCAGTTT CTAGAGAACC ATCAGATGTT TCCAGGGTGC CCCAAGGACC TGAAATGACC 4140
CTGTGCCTTA TTTGAACTAA CCAATCAGTT CGCTTCTCGC TTCTGTTCGC GCGCTTCTGC 4200

TCCCCGAGCT CAATAAAAGA GCCCACAACC CCTCACTCGG GGCGCCAGTC CTCCGATTGA 4260

CTGAGTCGCC CGG 4273

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CAAGAAAGCG GCTTCAGCTG TAAAGGACCT GGCCAGAATG TGGCTGTGAC CAGGGCACAC 60

CCTGACTCCC AAGGGAGGCG GCGGCGGCCT GAGCGGGGGG CCCGAGGAGG CCAGGTGTTT 120

TACAACAGCG AGTATGGGGA GCTGTCGGAG CCAAGCGAGG AGGACCACTG CTCCCCGTCT 180

GCCCGCGTGA CTTTCTTCAC AGACAACAGC TACTAA 216

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr Asn Thr Glu Ile Leu
1 5 10 15

Lys

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1836 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 168..1412

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GCGGCCGCGT CGACGCAAAA GTTGCGAGCC GCCGAGTCCC GGGAGACGCT CGCCCAGGGG 60

GGTCCCCGGG AGGAAACCAC GGGACTGGGA CCAGGAGAGG ACCTCAGCCT CACGCCCCAG 120

CCTGCGCCAG CCAACGGACC GGCCTCCCTG CTCCCGGTCC ATCCACC ATG CAC TTG 176
Met His Leu
1

CTG TGC TTC TTG TCT CTG GCG TGT TGC CTG CTC GCC GCT GCG CTG ATC 224
Leu Cys Phe Leu Ser Leu Ala Cys Ser Leu Leu Ala Ala Ala Leu Ile
5 10 15

CCC AGT CCG CGC GAG GCG CCC GCC ACC GTC GCC GCC TTC GAG TCG GGA 272
Pro Ser Pro Arg Glu Ala Pro Ala Thr Val Ala Ala Phe Glu Ser Gly
20 25 30 35

```
CTG GGC TTC TCG GAA GCG GAG CCC GAC GGG GGC GAG GTC AAG GCT TTT      320
Leu Gly Phe Ser Glu Ala Glu Pro Asp Gly Gly Glu Val Lys Ala Phe
                40                  45                  50

GAA GGC AAA GAC CTG GAG GAG CAG TTG CGG TCT GTG TCC AGC GTA GAT      368
Glu Gly Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser Ser Val Asp
            55                  60                  65

GAG CTG ATG TCT GTC CTG TAC CCA GAC TAC TGG AAA ATG TAC AAG TGC      416
Glu Leu Met Ser Val Leu Tyr Pro Asp Tyr Trp Lys Met Tyr Lys Cys
        70                  75                  80

CAG CTG CGG AAA GGC GGC TGG CAG CAG CCC ACC CTC AAT ACC AGG ACA      464
Gln Leu Arg Lys Gly Gly Trp Gln Gln Pro Thr Leu Asn Thr Arg Thr
    85                  90                  95

GGG GAC AGT GTA AAA TTT GCT GCT GCA CAT TAT AAC ACA GAG ATC CTG      512
Gly Asp Ser Val Lys Phe Ala Ala Ala His Tyr Asn Thr Glu Ile Leu
100                 105                 110                 115

AAA AGT ATT GAT AAT GAG TGG AGA AAG ACT CAA TGC ATG CCA CGT GAG      560
Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln Cys Met Pro Arg Glu
                120                 125                 130

GTG TGT ATA GAT GTG GGG AAG GAG TTT GGA GCA GCC ACA AAC ACC TTC      608
Val Cys Ile Asp Val Gly Lys Glu Phe Gly Ala Ala Thr Asn Thr Phe
            135                 140                 145

TTT AAA CCT CCA TGT GTG TCC GTC TAC AGA TGT GGG GGT TGC TGC AAC      656
Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys Gly Gly Cys Cys Asn
        150                 155                 160

AGG GAG GGG CTG CAG TGC ATG AAC ACC AGC ACA GGT TAC CTC AGC AAG      704
Arg Glu Gly Leu Gln Cys Met Asn Thr Ser Thr Gly Tyr Leu Ser Lys
    165                 170                 175

ACG TTG TTT GAA ATT ACA GTG CCT CTC TCA CAA GGC CCC AAA CCA GTC      752
Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln Gly Pro Lys Pro Val
180                 185                 190                 195

ACA ATC AGT TTT GCC AAT CAC ACT TCC TGC CGG TGC ATG TCT AAA CTG      800
Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg Cys Met Ser Lys Leu
                200                 205                 210

GAT GTT TAC AGA CAA GTT CAT TCA ATT ATT AGA CGT TCT CTG CCA GCA      848
Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg Arg Ser Leu Pro Ala
            215                 220                 225

ACA TTA CCA CAG TGT CAG GCA GCT AAC AAG ACA TGT CCA ACA AAC TAT      896
Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr Cys Pro Thr Asn Tyr
        230                 235                 240

GTG TGG AAT AAC TAC ATG TGC CGA TGC CTG GCT CAG CAG GAT TTT ATC      944
Val Trp Asn Asn Tyr Met Cys Arg Cys Leu Ala Gln Gln Asp Phe Ile
    245                 250                 255

TTT TAT TCA AAT GTT GAA GAT GAC TCA ACC AAT GGA TTC CAT GAT GTC      992
Phe Tyr Ser Asn Val Glu Asp Asp Ser Thr Asn Gly Phe His Asp Val
260                 265                 270                 275

TGT GGA CCC AAC AAG GAG CTG GAT GAA GAC ACC TGT CAG TGT GTC TGC     1040
Cys Gly Pro Asn Lys Glu Leu Asp Glu Asp Thr Cys Gln Cys Val Cys
                280                 285                 290

AAG GGG GGG CTT CGG CCA TCT AGT TGT GGA CCC CAC AAA GAA CTA GAT     1088
Lys Gly Gly Leu Arg Pro Ser Ser Cys Gly Pro His Lys Glu Leu Asp
            295                 300                 305

AGA GAC TCA TGT CAG TGT GTC TGT AAA AAC AAA CTT TTC CCT AAT TCA     1136
Arg Asp Ser Cys Gln Cys Val Cys Lys Asn Lys Leu Phe Pro Asn Ser
        310                 315                 320
```

```
TGT GGA GCC AAC AGG GAA TTT GAT GAG AAT ACA TGT CAG TGT GTA TGT      1184
Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr Cys Gln Cys Val Cys
    325                 330                 335

AAA AGA ACG TGT CCA AGA AAT CAG CCC CTG AAT CCT GGG AAA TGT GCC      1232
Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn Pro Gly Lys Cys Ala
340                 345                 350                 355

TGT GAA TGT ACA GAA AAC ACA CAG AAG TGC TTC CTT AAA GGG AAG AAG      1280
Cys Glu Cys Thr Glu Asn Thr Gln Lys Cys Phe Leu Lys Gly Lys Lys
            360                 365                 370

TTC CAC CAT CAA ACA TGC AGT TGT TAC AGA AGA CCG TGT GCG AAT CGA      1328
Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg Pro Cys Ala Asn Arg
        375                 380                 385

CTG AAG CAT TGT GAT CCA GGA CTG TCC TTT AGT GAA GAA GTA TGC CGC      1376
Leu Lys His Cys Asp Pro Gly Leu Ser Phe Ser Glu Glu Val Cys Arg
    390                 395                 400

TGT GTC CCA TCG TAT TGG AAA AGG CCA CAT CTG AAC TAAGATCATA          1422
Cys Val Pro Ser Tyr Trp Lys Arg Pro His Leu Asn
405                 410                 415

CCAGTTTTCA GTCAGTCACA GTCATTTACT CTCTTGAAGA CTGTTGGAAC AGCACTTAGC   1482

ACTGTCTATG CACAGAAAGA CTCTGTGGGA CCACATGGTA ACAGAGGCCC AAGTCTGTGT   1542

TTATTGAACC ATGTGGATTA CTGCGGGAGA GGACTGGCAC TCATGTGCAA AAAAAACCTC   1602

TTCAAAGACT GGTTTTCTGC CAGGGACCAG ACAGCTGAGG TTTTTCTCTT GTGATTTAAA   1662

AAAAGAATGA CTATATAATT TATTTCCACT AAAAATATTG TTCCTGCATT CATTTTTATA   1722

GCAATAACAA TTGGTAAAGC TCACTGTGAT CAGTATTTTT ATAACATGCA AAACTATGTT   1782

TAAAATAAAA TGAAAATTGT ATTATAAAAA AAAAAAAAAA AAAAAAAAAA GCTT         1836
```

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 415 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
Met His Leu Leu Cys Phe Leu Ser Leu Ala Cys Ser Leu Leu Ala Ala
 1               5                  10                  15

Ala Leu Ile Pro Ser Pro Arg Glu Ala Pro Ala Thr Val Ala Ala Phe
            20                  25                  30

Glu Ser Gly Leu Gly Phe Ser Glu Ala Glu Pro Asp Gly Gly Glu Val
        35                  40                  45

Lys Ala Phe Glu Gly Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser
    50                  55                  60

Ser Val Asp Glu Leu Met Ser Val Leu Tyr Pro Asp Tyr Trp Lys Met
65                  70                  75                  80

Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln Gln Pro Thr Leu Asn
                85                  90                  95

Thr Arg Thr Gly Asp Ser Val Lys Phe Ala Ala Ala His Tyr Asn Thr
            100                 105                 110

Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln Cys Met
        115                 120                 125
```

```
Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Ala Ala Thr
    130                 135                 140

Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys Gly Gly
145                 150                 155                 160

Cys Cys Asn Arg Glu Gly Leu Gln Cys Met Asn Thr Ser Thr Gly Tyr
                165                 170                 175

Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln Gly Pro
            180                 185                 190

Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg Cys Met
        195                 200                 205

Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg Arg Ser
    210                 215                 220

Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr Cys Pro
225                 230                 235                 240

Thr Asn Tyr Val Trp Asn Asn Tyr Met Cys Arg Cys Leu Ala Gln Gln
                245                 250                 255

Asp Phe Ile Phe Tyr Ser Asn Val Glu Asp Asp Ser Thr Asn Gly Phe
            260                 265                 270

His Asp Val Cys Gly Pro Asn Lys Glu Leu Asp Glu Asp Thr Cys Gln
        275                 280                 285

Cys Val Cys Lys Gly Gly Leu Arg Pro Ser Ser Cys Gly Pro His Lys
    290                 295                 300

Glu Leu Asp Arg Asp Ser Cys Gln Cys Val Cys Lys Asn Lys Leu Phe
305                 310                 315                 320

Pro Asn Ser Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr Cys Gln
                325                 330                 335

Lys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn Pro Gly
            340                 345                 350

Lys Cys Ala Cys Glu Cys Thr Glu Asn Thr Gln Lys Cys Phe Leu Lys
        355                 360                 365

Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg Pro Cys
    370                 375                 380

Ala Asn Arg Leu Lys His Cys Asp Pro Gly Leu Ser Phe Ser Glu Glu
385                 390                 395                 400

Val Cys Arg Cys Val Pro Ser Tyr Trp Lys Arg Pro His Leu Asn
                405                 410                 415
```

CLAIMS

What is claimed is:

1. A purified and isolated polypeptide which specifically binds to Flt4 receptor tyrosine kinase.

2. A purified and isolated polypeptide having the amino acid sequence of residues -102 to 317 shown in SEQ ID NO: 33.

3. A purified and isolated polypeptide having the amino acid sequence of residues 1 to 317 shown in SEQ ID NO: 33.

4. A fragment of the purified and isolated polypeptide according to claim 2, said fragment being capable of specifically binding to Flt4 receptor tyrosine kinase.

5. The fragment according to claim 4 having an apparent molecular weight of approximately 23 kD as assessed by SDS-PAGE under reducing conditions.

6. The fragment according to claim 4 comprising approximately amino acids 1-115 of SEQ ID NO: 33.

7. The fragment according to claim 4 comprising approximately amino acids 1-180 of SEQ ID NO: 33.

8. A purified and isolated polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, said polypeptide comprising a fragment of the purified and isolated polypeptide according to claim 2, said fragment being capable of specifically binding to Flt4 receptor tyrosine kinase.

9. A purified and isolated nucleic acid encoding the polypeptide according to claim 2.

10. A purified and isolated nucleic acid encoding the polypeptide according to claim 3.

11. The nucleic acid according to claim 9 having the sequence of nucleotides 352 to 1608 shown in SEQ ID NO: 32.

12. A purified and isolated nucleic acid encoding the fragment according to claim 5.

13. A purified and isolated nucleic acid encoding the fragment according to claim 6.

14. A purified and isolated nucleic acid encoding the fragment according to claim 7.

15. A nucleic acid comprising a VEGF-C encoding insert of plasmid pFLT4-L, deposited as ATCC accession Nu 97321.

16. A vector comprising the nucleic acid according to claim 9, 10, 12, or 13.

17. A vector comprising the nucleic acid according to claim 15, said vector being plasmid pFLT4-L, deposited as ATCC accession No. 97231.

18. A host cell transformed or transfected with the nucleic acid according to claim 9, 10, 12, or 13.

19. An antibody which is specifically reactive with an Flt4 receptor tyrosine kinase ligand.

20. An antibody which is specifically reactive with the polypeptide fragment according to claim 4.

21. An antibody of claim 20 which is a monoclonal antibody.

22. A pharmaceutical composition comprising a polypeptide fragment according to claim 4 in a pharmaceutically-acceptable diluent, adjuvant, or carrier.

23. A method of making a polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, said method comprising the steps of:

(a) expressing a nucleic acid according to claim 9, 10, 12, or 13 in a host cell; and

(b) purifying a polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase from said host cell or from a growth media of said host cell.

24. A polypeptide capable of specifically binding to Flt4 receptor tyrosine kinase, said polypeptide produced by the method according to claim 23.

25. A purified and isolated mammalian VEGF-C polypeptide.

26. A purified and isolated polynucleotide encoding the polypeptide of claim 25.

27. A purified and isolated polypeptide having the amino acid sequence of residues 1 to 415 of SEQ ID NO: 41.

28. A purified and isolated polypeptide capable of specifically binding to an Flt4 receptor tyrosine kinase, said polypeptide comprising a fragment of the purified and isolated polypeptide according to claim 27, said fragment being capable of specifically binding to said Flt4 receptor tyrosine kinase.

29. A purified and isolated nucleic acid encoding the polypeptide according to claim 27.

ABSTRACT

Provided are polypeptide ligands for the receptor tyrosine kinase, Flt4. Also provided are cDNAs and vectors encoding the ligands, pharmaceutical compositions and diagnostic reagents comprising the ligands, and methods of making and using the ligands.

Atty. Docket No: 28113/33118

MAIL ROOM

DECLARATION FOR PATENT APPLICATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that my residence, post office address and citizenship are as stated below next to my name; I believe that I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled "RECEPTOR LIGAND," the specification of which was filed on February 14, 1996, as Application Serial No. 08/601,132. I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by an amendment attached hereto. I acknowledge the duty to disclose to the Patent and Trademark Office all information known to me to be material to patentability as defined in 37 C.F.R. §1.56.

I hereby claim foreign priority benefits under 35 U.S.C. §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

(Application Serial Number)	(Country)	(Day/Month/Year Filed)	Priority Claimed Yes ☐	No ■
950624	Finland	13 February 1995		

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below:

(Application Serial Number)	(Day/Month/Year Filed)

I hereby claim the benefit under 35 U.S.C. §120 of any United States application(s) or PCT international application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) in the manner provided by the first paragraph of 35 U.S.C. §112, I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in 37 C.F.R. §1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

(Application Serial Number)	(Day/Month/Year Filed)	(Status-Patented, Pending or Abandoned)
08/340,011	14 November 1994	Pending
08/510,133	01 August 1995	Pending
08/585,895	17 January 1996	Pending

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. §1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: I hereby appoint as my attorneys, with full powers of substitution and revocation, to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

Alvin D. Shulman (19,412)	Trevor B. Jnike (25,542)	Richard A. Schnurr (30,890)	James J. Napoli (32,361)
Donald J. Breit (19,490)	Timothy J. Vezeau (26,348)	Anthony Nimmo (30,920)	Richard M. La Barge (32,254)
Owen J. Murray (22,111)	Carl E. Moore, Jr. (26,487)	Christine A. Dudzik (31,245)	Jeffry W. Smith (33,455)
Allen H. Gerstein (22,218)	Richard H. Anderson (26,526)	Kevin D. Hogg (31,839)	Douglass C. Hochstetler (33,710)
Nate F. Scarpelli (22,320)	Patrick D. Ertel (26,877)	Jeffrey S. Sharp (31,879)	Cynthia L. Schaller (34,245)
Edward M. O'Toole (22,477)	James P. Zeller (28,491)	Donald J. Pochopien (32,167)	Robert M. Gerstein (34,824)
Michael F. Borun (25,447)	William E. McCracken (30,195)	Martin J. Hirsch (32,237)	David A. Gass (38,153)

Send correspondence to: David A. Gass

FIRM NAME	PHONE NO.	STREET	CITY & STATE	ZIP CODE
Marshall, O'Toole, Gerstein, Murray & Borun	312-474-6300	6300 Sears Tower 233 South Wacker Drive	Chicago, Illinois	60606-6402

Full Name of First or Sole Inventor: Kari Alitalo	Citizenship: Finland
Residence Address - Street: Nyyrikintie 4A	Post Office Address - Street: Same
City (Zip): 02100 Espoo	City (Zip): Same
State or Country: FINLAND	State or Country: Same
Date: ■ Aug 6, 1996	Signature: ■

■ See second page for additional inventor

See reverse for relevant rules & statutes

Second Joint Inventor, if any	Citizenship
Vladimir Joukov	Russia
Residence Address - Street	Post Office Address - Street
Topeliuksenkatu 32G8	Same
City (Zip)	City (Zip)
00290 Helsinki FIX	Same
State or Country	State or Country
FINLAND	Same
Date	Signature
■ Aug. 6, 1996	■ V. Joukov

08601132

JOINT INVENTORS

"EXPRESS MAIL" mailing label No.
EF421126254US
Date of Deposit: February 14, 1996
I hereby certify that this paper (or fee) is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 CFR §1.10 on the date indicated above and is addressed to: Assistant Commissioner for Patents, Washington, D.C. 20231

David A. Gass

APPLICATION FOR UNITED STATES LETTERS PATENT

SPECIFICATION

TO ALL WHOM IT MAY CONCERN:

Be it known that we, **Kari Alitalo** a citizen of Finland, residing at Nyyrikintie 4A, 02100 Espoo, Finland, and **Vladimir Joukov** a citizen of Finland, residing at Topeliuksenkatu 32G8, 00290 Helsinki, Finland, have invented a new and useful **"RECEPTOR LIGAND"**, of which the following is a specification.

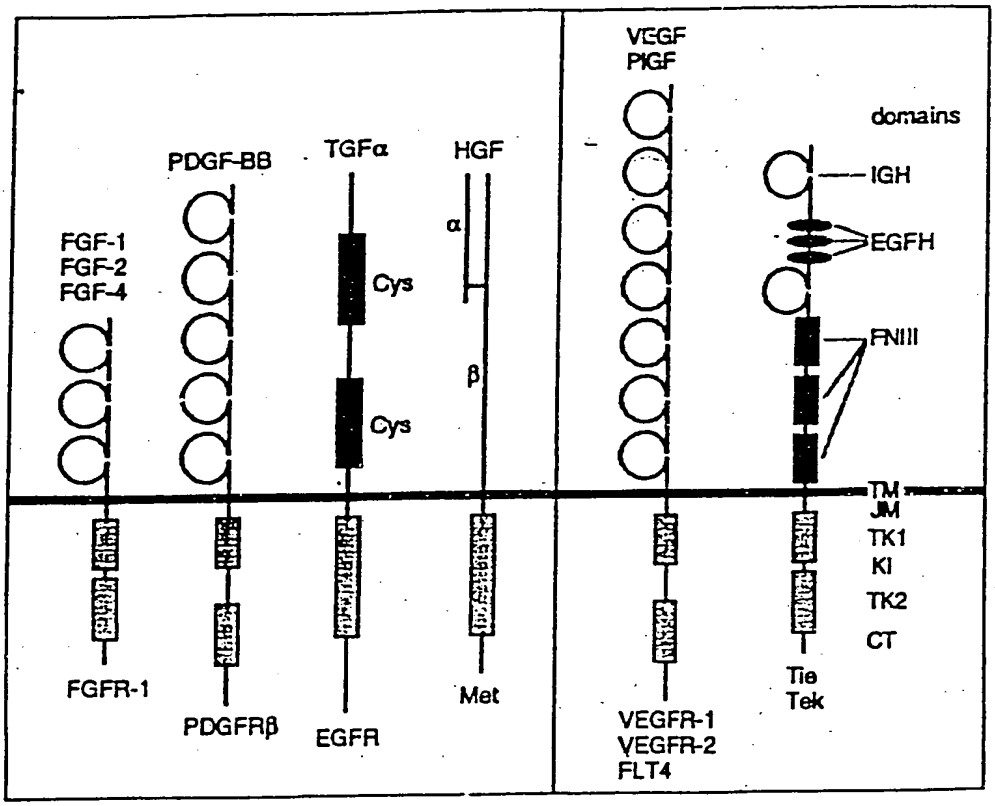

VEGF
PlGF
domains
IGH
EGFH
FNIII
PDGF-BB
TGFα
HGF
α
β
FGF-1
FGF-2
FGF-4
Cys
Cys
TM
JM
TK1
KI
TK2
CT
FGFR-1
PDGFRβ
EGFR
Met
Tie
Tek
VEGFR-1
VEGFR-2
FLT4

FIGURE 1

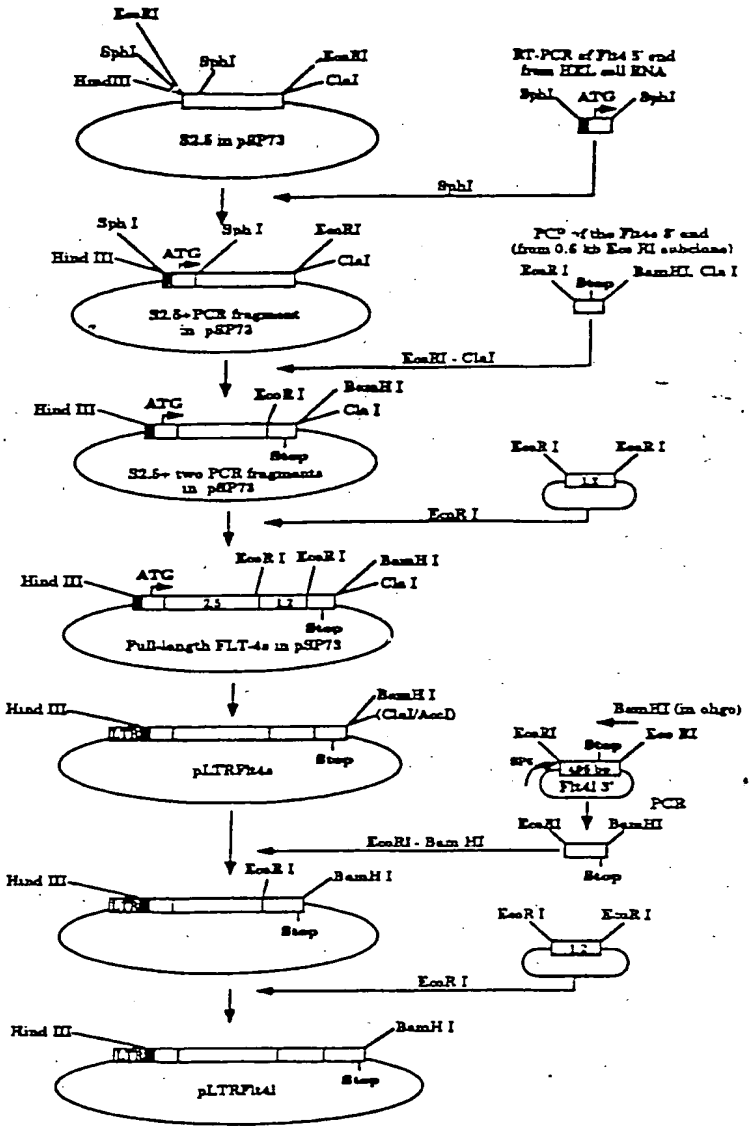
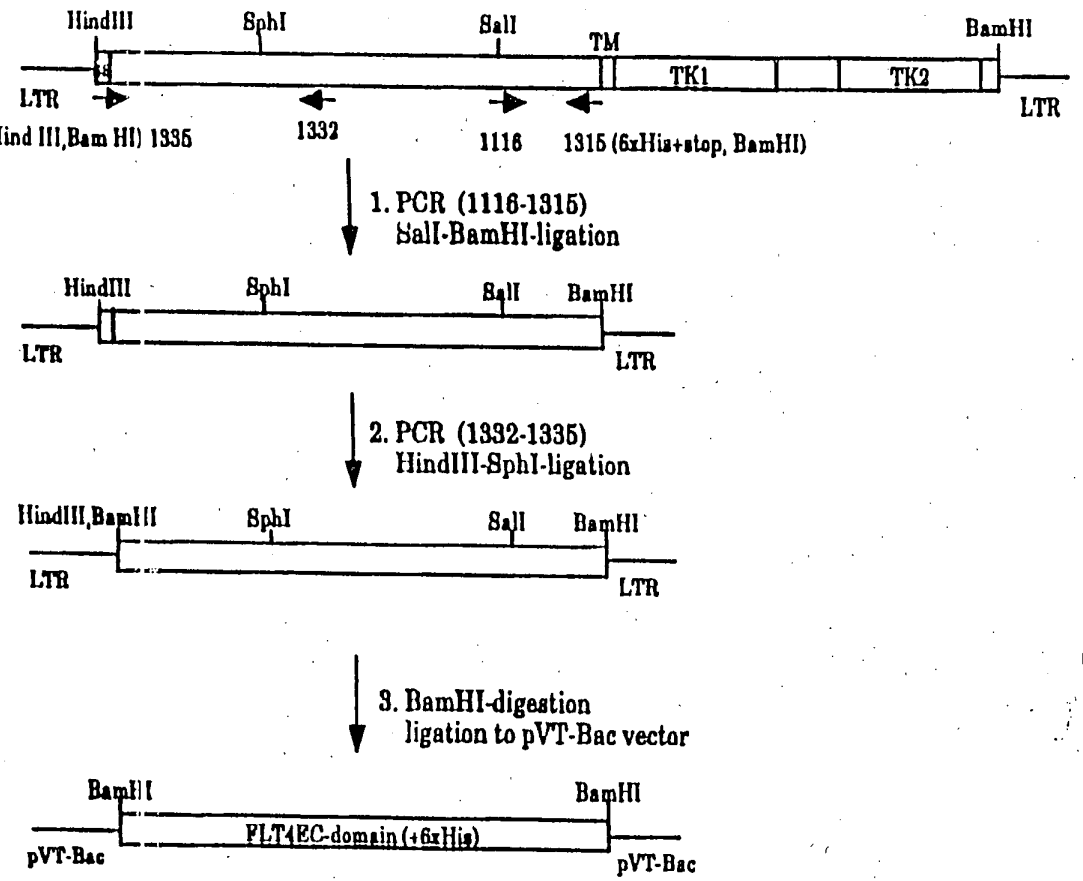

EcoRI
SphI
HindIII
ClaI
S2.6 in pSP73
RT-PCR of Flt4 5' end from HEL cell RNA
ATG
SphI - EcoRI
PCR of the Flt4s 5' end (from 0.6 kb Eco RI subclone)
BamHI
Stop
EcoRI - ClaI
S2.6+ two PCR fragments in pSP73
EcoR I
BamH I
Cla I
Hind III
Full-length FLT-4s in pSP73
pLTRFlt4s
(ClaI/AccI)
BamHI (in chgo)
Flt4l 5'
PCR
EcoRI - Bam HI
pLTRFlt4l

FIGURE 2

Figure 3



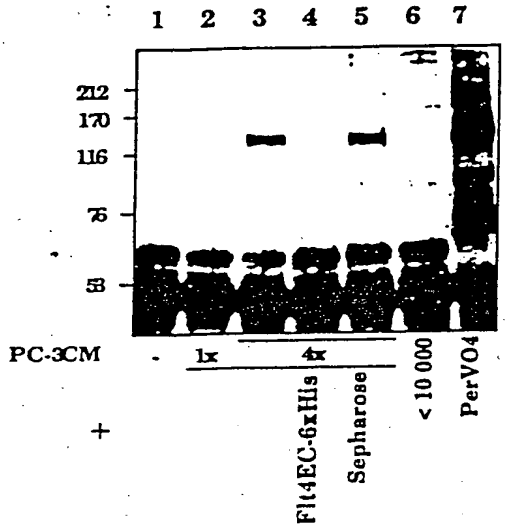

1 2 3 4 5 6 7
212
170
116
76
53
anti-PTyr
PC-3CM
-
+
1x
4x
Flt4EC-6xHis
Sepharose
< 10 000
PerVO4

FIGURE 4

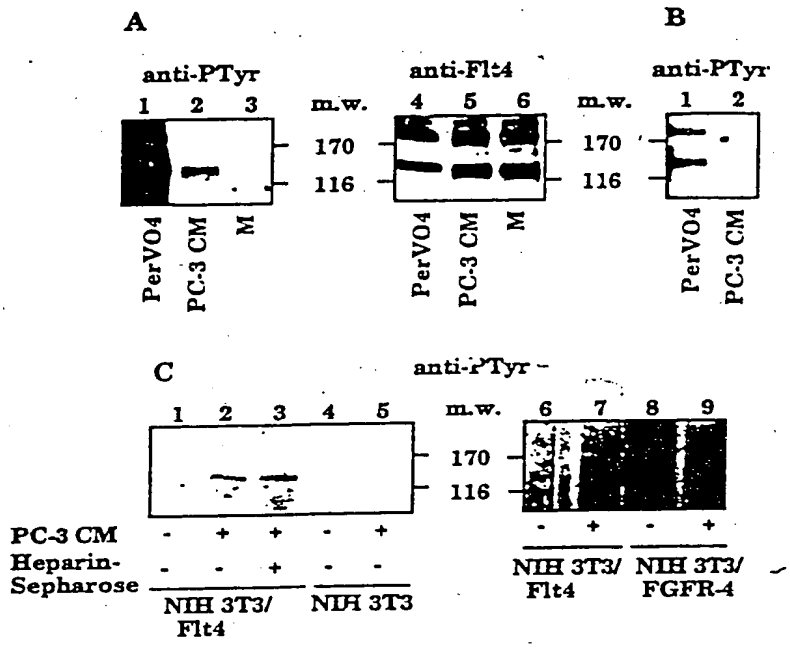

A
anti-PTyr
1 2 3
m.w.
170
116
PerVO4
PC-3 CM
M
anti-Flt4
4 5 6
m.w.
170
116
PerVO4
PC-3 CM
M
B
anti-PTyr
1 2
m.w.
170
116
PerVO4
PC-3 CM
C
anti-PTyr
1 2 3 4 5
m.w.
6 7 8 9
170
116
PC-3 CM - + + - +
Heparin-Sepharose - - + - -
NIH 3T3/ Flt4
NIH 3T3
- + - +
NIH 3T3/ Flt4
NIH 3T3/ FGFR-4

FIGURE 5

FIGURE 6

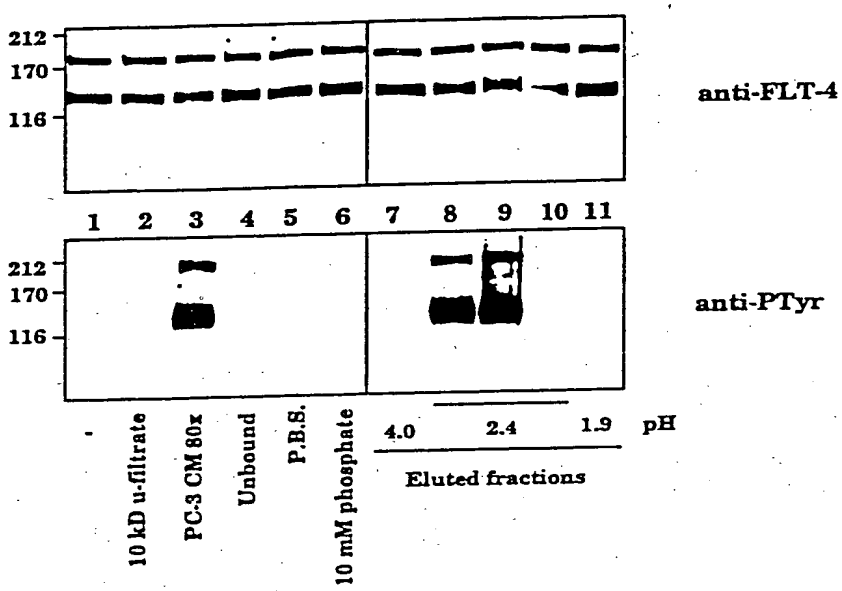

212
170
116
anti-FLT-4
1 2 3 4 5 6 7 8 9 10 11
212
170
116
anti-PTyr
-
10 kD u-filtrate
PC-3 CM 80x
Unbound
P.B.S.
10 mM phosphate
4.0
2.4
1.9
pH
Eluted fractions

FIGURE 7

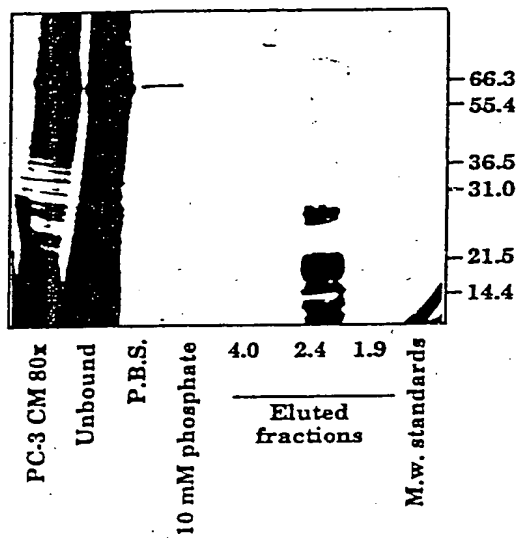

66.3
55.4
36.5
31.0
21.5
14.4
PC-3 CM 80x
Unbound
P.B.S.
10 mM phosphate
4.0
2.4
1.9
Eluted fractions
M.w. standards

FIGURE 8

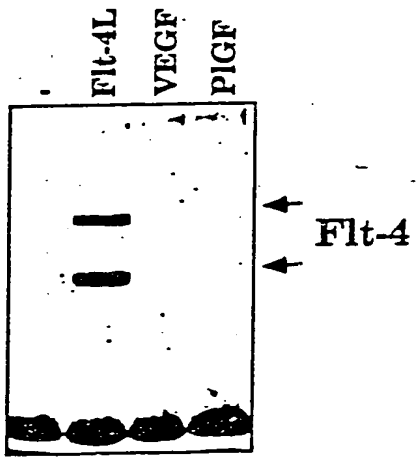
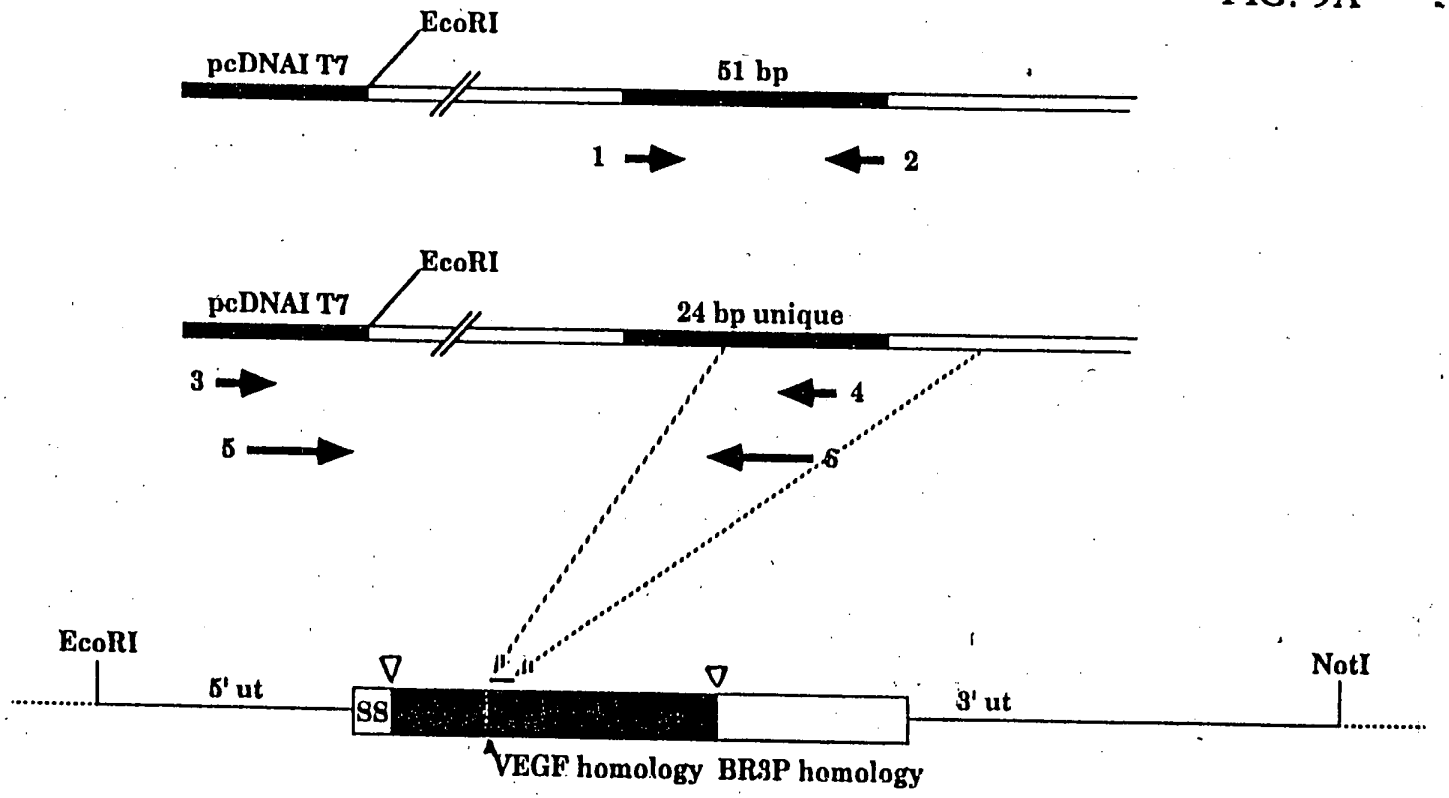


FIG. 9A



VEGF-C cDNA

```
        10                  30                  50
cccgccccgcctctccaaaaagctacaccgacgcggaccgcggcggcgtcctccctcgcc
        70                  90                 110
ctcgcttcacctcgcgggctccgaatgcggggagctcggatgtccggtttcctgtgaggc
       130                 150                 170
ttttacctgacacccgccgcctttccccggcactggctgggagggcgccctgcaaagttg
       190                 210                 230
ggaacgcggagccccggacccgctcccgccgcctccggctcgcccaggggggtcgccgg
       250                 270                 290
gaggagcccgggggagagggaccaggaggggcccgcggcctcgcaggggcgcccgcgccc
       310                 330                 350
ccacccctgccccgccagcggaccggtcccccaccccggtccttccaccatgcacttg
                                                  M  H  L
       370                 390                 410
ctgggcttcttctctgtggcgtgttctctgctcgccgctgcgctgctcccgggtcctcgc
L  G  F  F  S  V  A  C  S  L  L  A  A  A  L  L  P  G  P  R
       430                 450                 470
gaggcgcccgccgccgccgccgccttcgagtccggactcgacctctcggacgcggagccc
E  A  P  A  A  A  A  A  F  E  S  G  L  D  L  S  D  A  E  P
       490                 510                 530
gacgcgggcgaggccacggcttatgcaagcaaagatctggaggagcagttacggtctgtg
D  A  G  E  A  T  A  Y  A  S  K  D  L  E  E  Q  L  R  S  V
       550                 570                 590
tccagtgtagatgaaotcatgaotgtaotctaoccagaatattggaaaatgtacaagtgt
S  S  V  D  E  L  M  T  V  L  Y  P  E  Y  W  K  M  Y  K  C
       610                 630                 650
cagctaaggaaaggaggctggcaacataacagagaacaggccaacctcaactcaaggaca
Q  L  R  K  G  G  W  Q  H  N  R  E  Q  A  N  L  N  S  R  T
```

FIG. 9B (1 OF 3)

```
      670                 690                 710
gaagagactataaaatttgctgcagcacattataatacagagatcttgaaaagtattgat
E  E  T  I  K  F  A  A  A  H  Y  N  T  E  I  L  K  S  I  D
      730                 750                 770
aatgagtggagaaagactcaatgcatgccacgggaggtgtgtatagatgtggggaaggag
N  E  W  R  K  T  Q  C  M  P  R  E  V  C  I  D  V  G  K  E
      790                 810                 830
tttggagtcgcgacaaacaccttctttaaacctccatgtgtgtccgtctacagatgtggg
F  G  V  A  T  N  T  F  F  K  P  P  C  V  S  V  Y  R  C  G
      850                 870                 890
ggttgctgcaatagtgaggggctgcagtgcatgaacaccagcacgagctacctcagcaag
G  C  C  N  S  E  G  L  Q  C  M  N  T  S  T  S  Y  L  S  K
      910                 930                 950
acgttatttgaaattacagtgcctctctctcaaggccccaaaccagtaacaatcagtttt
T  L  F  E  I  T  V  P  L  S  Q  G  P  K  P  V  T  I  S  F
      970                 990                1010
gccaatcacacttcctgccgatgcatgtctaaactggatgtttacagacaagttcattcc
A  N  H  T  S  C  R  C  M  S  K  L  D  V  Y  R  Q  V  H  S
     1030                1050                1070
attattagacgttccctgccagcaacactaccacagtgtcaggcagcgaacaagacctgc
I  I  R  R  S  L  P  A  T  L  P  Q  C  Q  A  A  N  K  T  C
     1090                1110                1130
cccaccaattacatgtggaataatcacatctgcagatgcctggctcaggaagattttatg
P  T  N  Y  M  W  N  N  H  I  C  R  C  L  A  Q  E  D  F  M
     1150                1170                1190
ttttcctcggatgctggagatgactcaacagatggattccatgacatctgtggaccaaac
F  S  S  D  A  G  D  D  S  T  D  G  F  H  D  I  C  G  P  N
     1210                1230                1250
aaggagctggatgaagagacctgtcagtgtgtctgcagagcggggcttcggcctgccagc
K  E  L  D  E  E  T  C  Q  C  V  C  R  A  G  L  R  P  A  S
     1270                1290                1310
tgtggaccccacaaagaactagacagaaaactcatgccagtgtgtctgtaaaaacaaactc
C  G  P  H  K  E  L  D  R  N  S  C  Q  C  V  C  K  N  K  L
```

FIG. 9B (2 0F 3)

```
      1330                  1350                  1370
ttccccagccaatgtggggccaaccgagaatttgatgaaaacacatgccagtgtgtatgt
F  P  S  Q  C  G  A  N  R  E  F  D  E  N  T  C  Q  C  V  C
      1390                  1410                  1430
aaaagaacctgccccagaaatcaacccctaaatcctggaaaatgtgcctgtgaatgtaca
K  R  T  C  P  R  N  Q  P  L  N  P  G  K  C  A  C  E  C  T
      1450                  1470                  1490
gaaagtccacagaaatgcttgttaaaaggaaagaagttccaccaccaaacatgcagctgt
E  S  P  Q  K  C  L  L  K  G  K  K  F  H  H  Q  T  C  S  C
      1510                  1530                  1550
tacagacggccatgtacgaaccgccagaaggcttgtgagccaggattttcatatagtgaa
Y  R  R  P  C  T  N  R  Q  K  A  C  E  P  G  F  S  Y  S  E
      1570                  1590                  1610
gaagtgtgtcgttgtgtcccttcatattggaaaagaccacaaatgagctaagattgtact
E  V  C  R  C  V  P  S  Y  W  K  R  P  Q  M  S
      1630                  1650                  1670
gttttccagttcatcgattttctattatggaaaactgtgttgccacagtagaactgtctg
      1690                  1710                  1730
tgaacagagagacccttgtgggtccatgctaacaaagacaaaagtctgtctttcctgaac
      1750                  1770                  1790
catgtggataactttacagaaatggactggagctcatctgcaaaaggcctcttgtaaaga
      1810                  1830                  1850
ctggttttctgccaatgaccaaacagccaagattttcctcttgtgatttctttaaaagaa
      1870                  1890                  1910
tgactatataatttatttccactaaaaatattgtttctgcattcatttttatagcaacaa
      1930                  1950                  1970
caattggtaaaactcactgtgatcaatatttttatatcatgcaaaatatgtttaaaataa
      1990
aatgaaaattgtattat
```

FIG. 9B (3 0F 3)

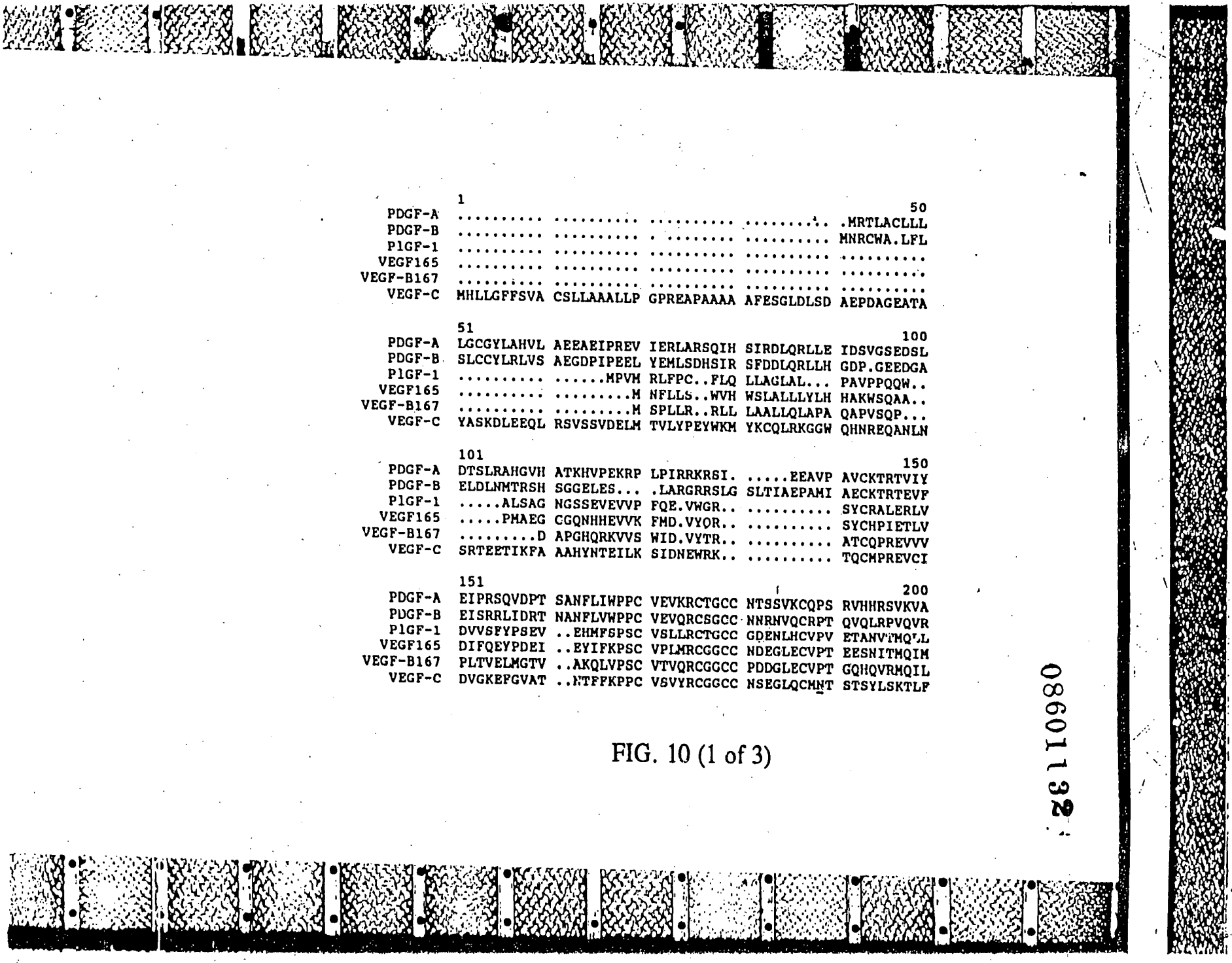


FIG. 10 (1 of 3)

```
      201                                                                250
PDGF-A     KVEYVRKKPK LKEVQVRLEE HLECACAT.. .......... TSLNPDYREE
PDGF-B     KIEIVRKKPI FKKATVTLED HLACKCETVA AARPVTRSPG GSQEQRAKTP
PlGF-1     KIRSG..DRP .SYVELTFSQ HVRCECRPLR EK........ ..........
VEGF165    RIKPH..QGQ .HIGEMSFLQ HNKCECRPKK DR........ ..........
VEGF-B167  MIRYP..SSQ ..LGEMSLEE HSQCECRPKK KD........ ..........
VEGF-C     EITVPLSQGP .KPVTISFAN HTSCRCMSKL DVYRQVHSII RRSLPATLPQ

      251                                                          300
PDGF-A     DTDVR..... .......... .......... .......... ..........
PDGF-B     QTRVTIRTVR VRRPPKGKHR KFKHTHDKIA LKETLGA... ..........
PlGF-1     .......... .......... .......... .MKPERCGDA VPRR......
VEGF165    .......... .......... .......... .ARQENPCGP CSERRKHLFV
VEGF-B167  .......... .......... .........S AVKPDSPRPL CPRCTQHHQR
VEGF-C     CQAANKTCPT NYMWNNHICR CLAQEDFMFS SDAGDDSTDG FHDICGPNKE

      301                                                          350
PDGF-A     .......... .......... .......... .......... ..........
PDGF-B     .......... .......... .......... .......... ..........
PlGF-1     .......... .......... .......... .......... ..........
VEGF165    QDPQTCKCSC KNTDS.RCKA RQLELNERTC RCDKPRR... ..........
VEGF-B167  PDPRTCRCRC RRRSFLRCQG RGLELNPDTC RCRKLRR... ..........
VEGF-C     LDEETCQCVC RAGLRPASCG PHKELDRNSC QCVCKNKLFP SQCGANREFD
```

FIG. 10 (2 of 3)

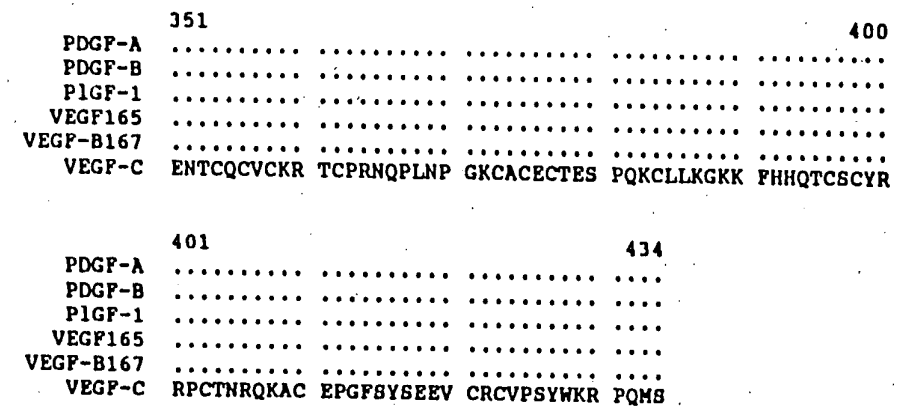


FIG. 10 (3 of 3)

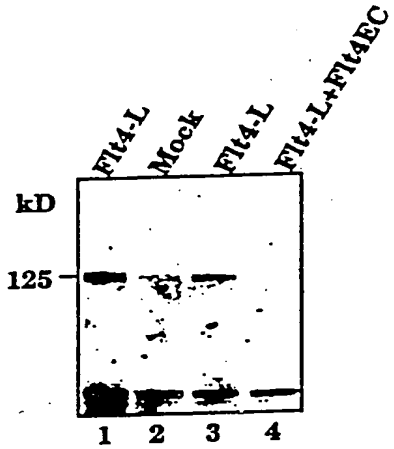


FIGURE 11

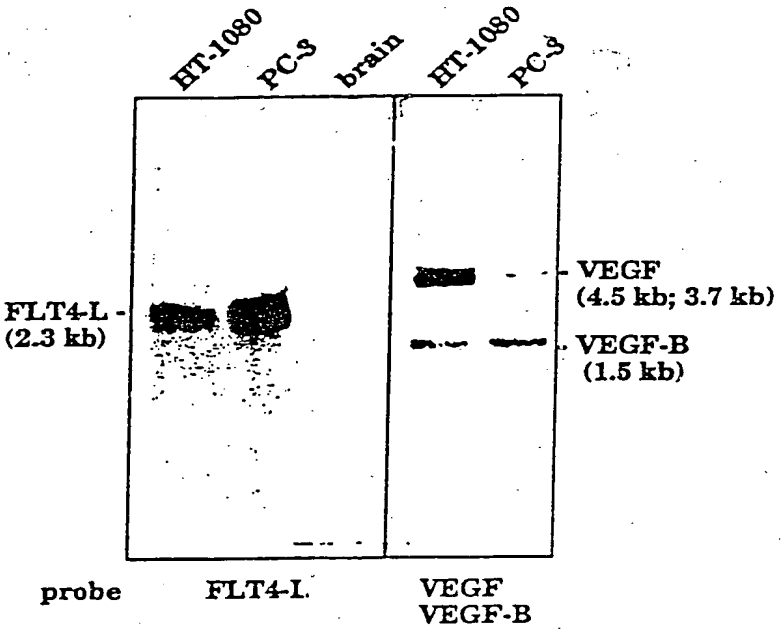

HT-1080
PC-3
brain
HT-1080
PC-3
FLT4-L
(2.3 kb)
VEGF
(4.5 kb; 3.7 kb)
VEGF-B
(1.5 kb)
probe
FLT4-L
VEGF
VEGF-B

FIGURE 12

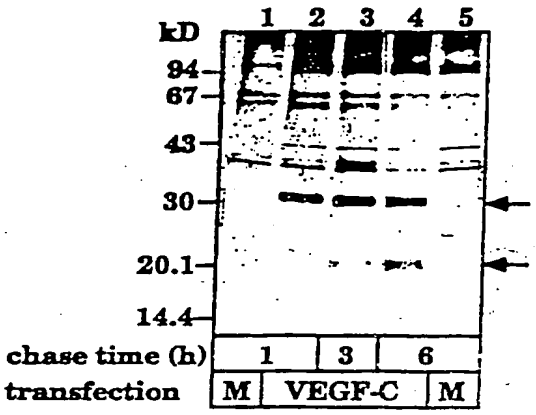

kD
1
2
3
4
5
94
67
43
30
20.1
14.4
chase time (h)
1
3
6
transfection
M
VEGF-C
M

FIG. 13A

FIG. 13B

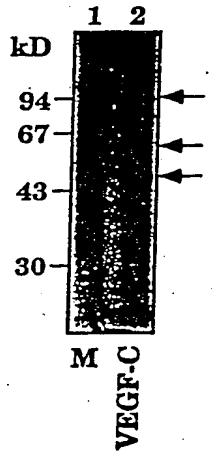


FIG. 14A

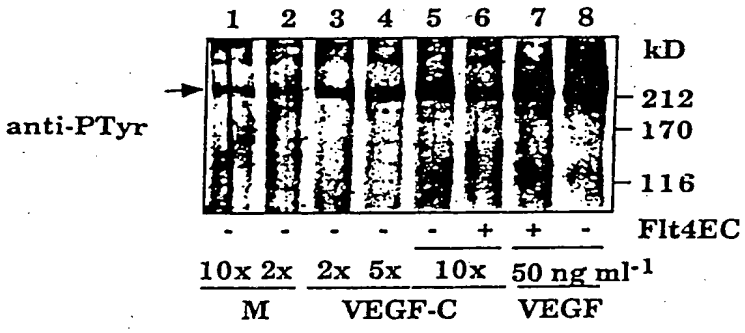

1 2 3 4 5 6 7 8
kD
anti-PTyr
212
170
116
- - - - - + + -
Flt4EC
10x 2x 2x 5x 10x 50 ng ml-1
M
VEGF-C
VEGF

FIG. 14B

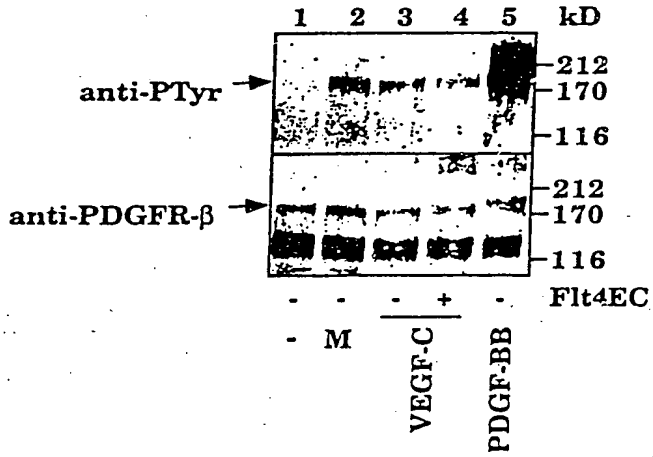

1 2 3 4 5 kD
anti-PTyr
212
170
116
anti-PDGFR-β
212
170
116
- - - + - Flt4EC
- M
VEGF-C
PDGF-BB

FIG. 15A

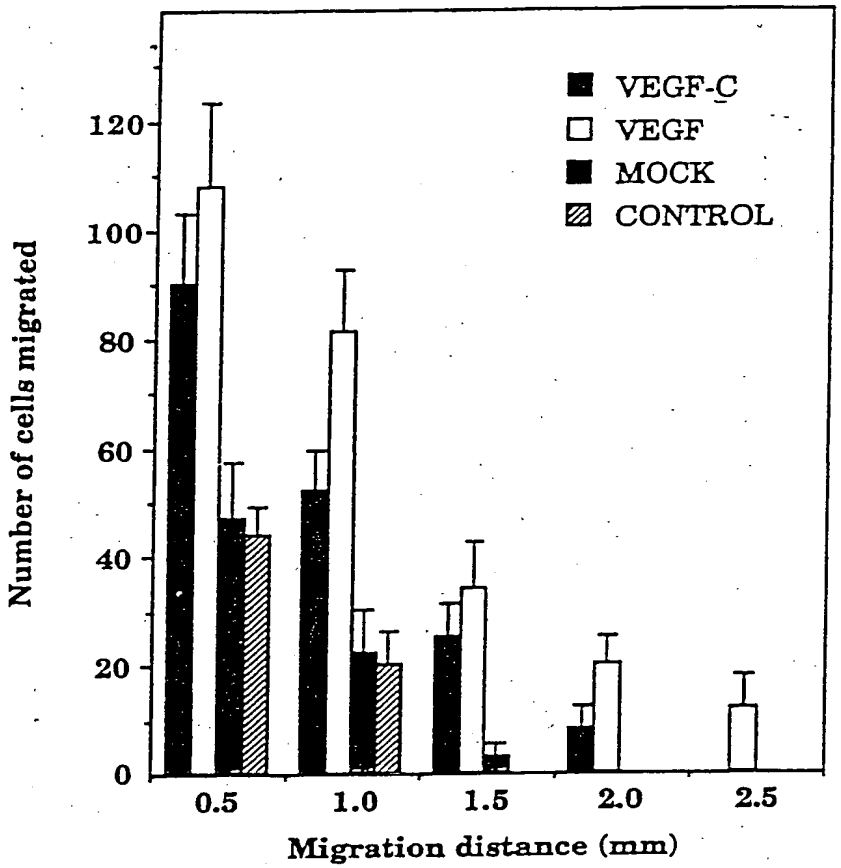

VEGF-C
VEGF
MOCK
CONTROL
120
100
80
60
40
20
0
Number of cells migrated
0.5
1.0
1.5
2.0
2.5
Migration distance (mm)

FIG. 15B

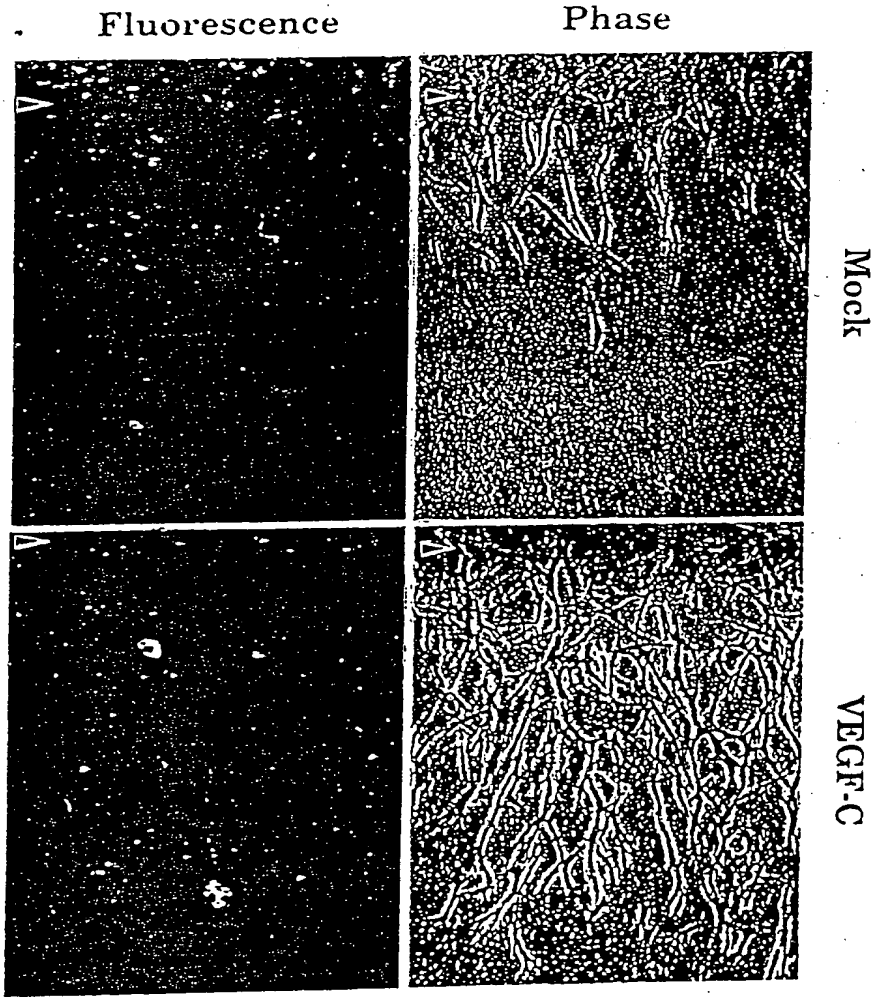
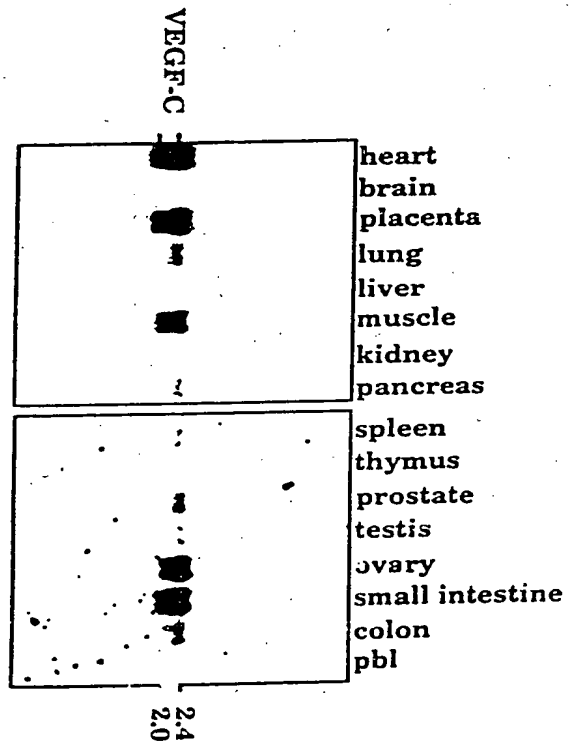


FIG. 16A


VEGF-C
heart
brain
placenta
lung
liver
muscle
kidney
pancreas
spleen
thymus
prostate
testis
ovary
small intestine
colon
pbl
2.4
2.0

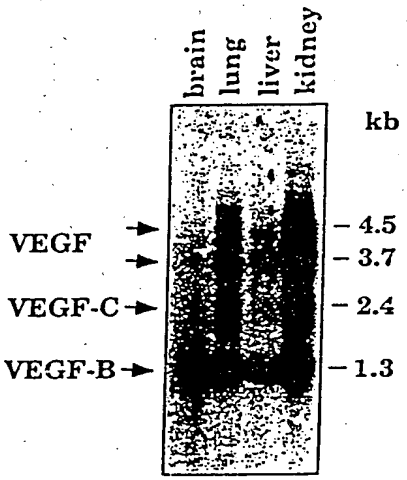
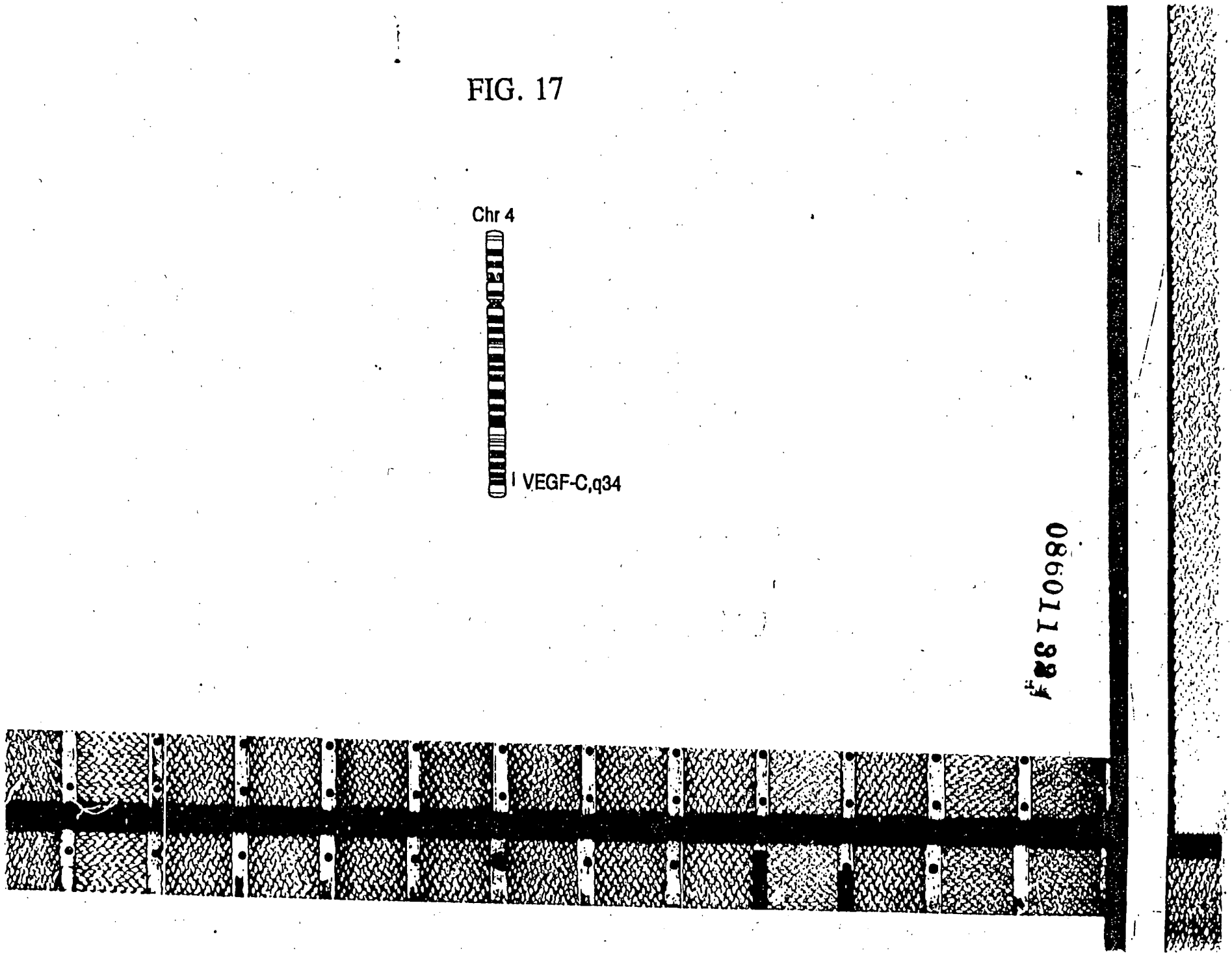


FIG. 16B

FIG. 17



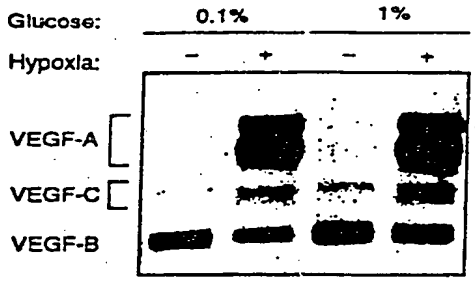


FIG. 18